

MILLIKANS BIOSEMANTIK UND DAS PROBLEM DER
MISSREPRÄSENTATION

Abhandlung
zur Erlangung der Doktorwürde
der Philosophischen Fakultät
der
Universität Zürich

vorgelegt von

André Wunder

Angenommen im Herbstsemester 2015
auf Antrag der Promotionskommission:

Prof. Dr. Hans-Johann Glock «(hauptverantwortliche
Betreuungsperson)»
Prof. Dr. Markus Wild

Zürich, 2015

Für meine Eltern, denen ich soviel verdanke

Danksagung

Für konstruktive Diskussionen und Kommentare, die wesentlich zur Entstehung dieses Buches beitragen haben, danke ich Hanjo Glock, Ruth Millikan, Tony Pacyna, Anne-Katrin Schlegel, Ulrich Stegmann, Christian Steiner, Christian Tewes und Markus Wild.

Ganz besonders möchte ich aber Hanno Birken-Bertsch für seine wertvollen Ratschläge und hilfreichen Hinweise danken, auf die ich nun schon seit meinen Studienanfängen vertrauen darf.

Ebenso gilt besonderer Dank meinem Schulfreund, Michael Friedrich, der durch seine Wissbegier und seinen Forscherdrang für mich stets eine Quelle der Inspiration und Motivation gewesen ist.

Schliesslich möchte ich auch meiner Lebensgefährtin, Sybille Pohl, dafür danken, dass sie mich im Zustand misstrautigen Grübelns und Zweifelns so tapfer ertragen hat.

Der grösste Dank gebührt aber meinen Eltern, denen ich dieses Buch widme, weil sie immer an mich geglaubt haben und mich bei allen, was ich gemacht habe, mit Rat und Tat unterstützt haben.

Inhalt

Einleitung	6
1. Kontext und Rahmen der Untersuchung	7
1.1. Kontext: Desiderat beim Informationsbegriff	8
1.1.1. Informationsbegriff in Kognitionswissenschaften	8
1.1.2. Kritik am Informationsbegriff	10
1.1.3. Vertiefung der Kritik: Der Homunkuluseinwand	11
1.1.4. Naturalistische Theorie des Gehalts als Ausweg	13
1.2. Naturalistische Theorie des Gehalts	15
1.2.1. Das Ziel naturalistischer Theorien des Gehalts	15
1.2.2. Theorieverständnis	16
1.2.3. Naturalismusverständnis	17
1.2.4. Gehaltverständnis	19
1.2.5. Bedingung: Fehler müssen möglich sein	24
1.3. Das Problem der Missrepräsentation	26
1.3.1. Das Problem der Missrepräsentation in naturalistischen Ansätzen	26
1.3.2. Disjunktionsproblem	27
1.3.3. Missrepräsentationsverständnis	28
1.3.4. Der teleosemantische Ansatz	30
1.3.4.1. Die grundlegende Idee anhand von Dretske	30
1.3.4.2. Herausforderung: Unbestimmtheit der Funktion	32
1.3.4.3. Die Einschätzung der teleosemantischen Lösung	34
1.3.5. Anspruch an die Lösung des Problems	36
1.4. Ziel der Arbeit	39
1.4.1. Millikans Lösung überprüfen	39
1.4.2. Warum Millikans Lösung	41
1.4.3. Millikans Anspruch	43
1.5. Methode	45
2. Darstellung von Millikans Lösung	47
2.1. Millikans Theorie des Gehalts	48
2.1.1. Millikans Zeichenbegriff	48
2.1.1.1. Millikans Zeichenbegriff: Die Grundidee	48
2.1.1.2. Millikans Definition eines Zeichens	50
2.1.1.3. Erläuterung technischer Ausdrücke	53
2.1.1.3.1. Reproduktion	53
2.1.1.3.2. REF erster Ordnung	57

2.1.1.3.3. Eigenfunktion	60
2.1.1.3.4. Normale Erklärung, Normale Bedingung	73
2.1.1.3.5. REF höherer Ordnung	76
2.1.1.3.6. Relationale EF	85
2.1.1.3.7. Standardisierte EF	91
2.1.1.3.8. Mapping Function	98
2.1.2. Gehaltsbestimmung bei Millikan	102
2.2. Fehlrepräsentation bei Millikan	109
2.2.1. Zwei Lösungsansätze	109
2.2.2. ab Normale Bedingungen	112
2.2.2.1. Umweltbed. Fehlfunktion	114
2.2.2.2. Umweltbed. Fehlrepräsentation	118
2.2.3. Missgebildete Funktionsträger	121
2.2.3.1. Missbildungsbed. Fehlfunktion	121
2.2.3.2. Missbildungsbed. Fehlrepräsentation	124
3. Bewertung von Millikans Lösung	129
3.1. Vorgehen	130
3.2. Einzelne Einwände	131
3.2.1. Fodor	131
3.2.2. Pietroski	146
3.2.3. Keil	159
3.2.4. Davies	165
3.2.5. Perlmann	176
3.2.6. Mendola	182
3.2.7. Stegmann	187
3.2.8. Nanay	194
3.2.9. Wunder	204
Fazit	215
Literatur	216
CV	231

Einleitung

Das zentrale Ziel dieser Arbeit ist es die Frage zu beantworten, ob Millikans Gehaltstheorie, die Biosemantik, das Problem der Missrepräsentation lösen kann oder ob ihr dies nicht gelingt.

Die Antwort wird insgesamt positiv ausfallen: Millikans Biosemantik enthält tatsächlich eine Lösung für das Problem der Missrepräsentation. Diese Lösung ist allerdings, wie sich herausstellen wird, in bedeutsamer Weise eingeschränkt.

Die Arbeit gliedert sich in drei Teile.

Im *ersten Teil* wird zunächst der Kontext erläutert, in dem die Beantwortung der Frage meines Erachtens von Bedeutung ist. Der Kontext ist das Informationsverarbeitungsparadigma der Kognitionswissenschaften und der damit verbundene Homunkuluseinwand. Eine naturalistische Gehaltstheorie, die das Problem der Missrepräsentation löst, bietet den Kognitionswissenschaften einen Weg an der Auffassung von Kognition als Informationsverarbeitung festzuhalten und trotzdem dem Homunkuluseinwand zu entgehen. Im Anschluss an die Erläuterung des Kontexts wird der Rahmen der Untersuchung festgelegt. In diesem Zusammenhang wird geklärt, welches Verständnis von Theorie, Naturalismus und Gehalt diese Untersuchung voraussetzt, worin das Problem der Missrepräsentation eigentlich besteht, wie eine Lösung aussehen könnte und welchen Ansprüchen sie genügen sollte; zudem wird die Wahl von Millikans Theorie für die Untersuchung begründet und das methodische Vorgehen der Untersuchung angegeben.

Der *zweite Teil* widmet sich voll und ganz der Darstellung von Millikans Lösung des Problems der Missrepräsentation, deren genaues Verständnis die Grundlage für die Beantwortung der Untersuchungsfrage ist. Ausgehend von Millikans Zeichenbegriff werden zunächst alle wichtigen technischen Begriffe von Millikan eingehend erläutert. Darauf aufbauend wird dargestellt, wie nach Millikans Auffassung der Gehalt eines Gehaltsträgers zu bestimmen ist. Aus diesen theoretischen Grundlagen von Millikans Theorie ergeben sich dann nicht nur ein, sondern zwei weitestgehend unabhängige Lösungsansätze für das Problem der Missrepräsentation. Beide Ansätze werden ausführlich dargestellt.

Der *dritte Teil* ist die kritische Prüfung von Millikans Lösungsansätzen. Die Prüfung erfolgt anhand von acht Einwänden gegen Millikans Lösung. Die Behandlung jedes Einwandes umfasst dessen Rekonstruktion und ausführliche Diskussion, zudem wird ein eigener Einwand präsentiert werden.

Abschliessend werden die Ergebnisse der Untersuchung in einem Fazit zusammengefasst.

Teil 1:

**Kontext und Rahmen der
Untersuchung**

1.1. Kontext: Ein Desiderat beim Informationsbegriff

1.1.1. Der Informationsbegriff in den Kognitionswissenschaften

Der Informationsbegriff ist einer der disziplinenübergreifenden Begriffe in den Naturwissenschaften. Er spielt eine bedeutende Rolle in der Physik, der Biologie und den Kognitionswissenschaften (Vgl. Adriaans 2012; Hiley 2004, 199; Lyre 2002, 8 und Janich, 1999, 9).

„The concept of information is a central unifying concept in the sciences. It plays crucial roles in physics, computation and control theory, biology, cognitive neuroscience, and of course the social sciences.“ (Deacon 2010, 150)

Für die Kognitionswissenschaften ist der Informationsbegriff von herausragender Bedeutung, weil er für ihre Auffassung von Kognition, ihr Kognitionsparadigma, zentral ist. Kognition wird in diesen Wissenschaften nahezu ausnahmslos als Informationsverarbeitung verstanden (Vgl. Lyre 2002, 125; Bechtel et al 1999, 91; Bischof 2009, 493; Haughland 1998, 26; Stout 2006, 44; Rakover 2007, 178; Noe 2009, 160; Melser 2004, 2)¹. Kognitive Prozesse sind Prozesse der Informationsverarbeitung.

„The broad movement which superseded behaviourism, and which has, *to date* [A.W.], proved far more theoretically progressive, is cognitivism. Cognitive psychology treats human brains and the brains of other intelligent organisms – as, at bottom, information-processing systems.“ (Botterill and Carruthers 2003, 17)

Besonderes Interesse gilt dabei dem Fluss der Information von den Sinnesorganen zum Gehirn und der anschließenden Verarbeitung der Information, die dann die Grundlage für Handeln und Gedächtnis bildet (Vgl. Kandel et al 2011, 329; Bermudez 2005, 208). Das so verstandene kognitive System, d.h. das Gehirn oder einzelne Neuronenverbände, verschlüsseln Informationen, leiten diese weiter und speichern diese (Vgl. exemplarisch Stein et al 1998, 433; Dennett 1996, 68f; Pellen 2008, 441; Frith 2007, 92; Goldblum 2004, 13;

¹ Eine wichtige Gegenposition ist die sogenannte Dynamical Systems Theory (DST), die am prominentesten von van Gelder (Van Gelder 1995, 1998; Van Gelder & Port 1995) vertreten wird. Es ist aber fraglich, ob die DST eine echte Konkurrenz darstellt, da sie nach Meinung ihrer Kritiker kein Kognitionsparadigma anzubieten hat. „Van Gelder misunderstands the former [cognitivism, A.W.]. Cognitivism is more than just the hypothesis that computer science supplies the appropriate modelling tools for cognitive science it involves a strong positive conception of the nature of cognitive agents. Cognitivism explains epistemic capacity (the system possesses internal symbolical states which representationally encode information about the environment), rationality (in terms of algorithms that optimise for truth preservation when computing belief formation, and for desire satisfaction when computing goal-directed output), and adaptive capacity (in terms of the rational response to the represented state of the environment). Regardless of the ultimate adequacy of these explanations, the status of cognitivism as a cognitive paradigm hinges on the fact that it provides a systematic conception of the nature of cognition. In contrast, DST has no rival cognitive paradigm to offer, only a general recommendation to pay attention to dynamical modelling when studying cognitive phenomena.“ (Christensen and Hooker 2004, 45).

Eine weitere Alternative kann in der Auffassung von *embodied cognition* gesehen werden (Vgl. z.B. Clark 2008), insbesondere wenn sie eine radikale Ausprägung annimmt (Vgl. Chemero 2009).

Grossenbacher 2001, 12; Mausfeld 2005, 18 [65], Uhlhaas and Singer 2011, 3; McDermott 2001, 10).

„The brain is a storehouse of information brought in through different sensory pathways, and acquired at different times and places. Information contained in occurrent sensory states that has a fixed bearing on behavior is used by sensorimotor subsystems without conscious registration.“ (Ismael 2007, 204)

Dieses Kognitionsverständnis setzt dabei das Vorhandensein von wie auch immer gearteten *Trägern semantischer Information* innerhalb des kognitiven Systems voraus (Vgl. Bechtel 2008, ix; Gallistel 2003, 228; Thornton 2006, 125).

„According to the received view in cognitive science, in order to explain cognitive abilities we need to appeal to information-bearing states inside the system. Such states, by virtue of the semantic information they carry about the world, qualify as representations. Cognitivists conceive of these representations as symbols in a computational "language of thought," and connectionists as constrained patterns of network activity corresponding to phase space "attractors" (regions of phase space toward which all nearby trajectories converge). In either case there is a strong tendency to adopt an objectivist conception of representation: representations are internal structures that encode context-independent information about the world, and cognition is the processing of such information.“ (Thompson 2007, 52)

Die Auffassung von Kognition als Informationsverarbeitung setzte sich in den 1960ern durch und verdrängte damit den Behaviorismus als vorherrschende Doktrin in der Psychologie (Vgl. Hunt 2005, 3; Heil 2004, 52). Dieses Geschehen wird in der Literatur oft in Anlehnung an Kuhn als „cognitive revolution“ bezeichnet (Vgl. Bechtel et al 1998, 6; Frith 2007, 112-116). Maßgeblich für diese Entwicklung waren wichtige Arbeiten auf dem Gebiet der Ingenieurwissenschaften, insbesondere der Nachrichtentechnik (Vgl. Velmans 2009, 65-66). Allen voran ist hier die Informationstheorie² von Shannon und Weaver zu nennen (Vgl. Bischof 2009 117, Garner 1988, 21)³. Die Aufnahme der Grundideen und -begriffe dieser Theorie in die Psychologie erfolgte in einem rasanten Tempo (Vgl. Garner 1988, 22).

„Psychology must have been waiting, even eagerly, for such a theory [information theory, A.W.], because it was introduced to the field almost immediately after Shannon's publication.“ (Garner 1988, 21)

Diese schnelle Aufnahme hängt wohl auch damit zusammen, dass der rein mathematisch formulierte Informationsbegriff von Shannon jeglichem Metaphysikverdacht standzuhalten scheint. Der Informationsbegriff bot daher die Möglichkeit sich wieder den inneren mentalen Prozessen zuzuwenden, ohne dass man von den Behavioristen der Unwissenschaftlichkeit besichtigt werden konnte (Bechtel et al 1998, 6; Bischof 2009, 496).

² Eine kurze und verständliche Darstellung zur nachrichtentechnische Informationstheorie findet sich z.B. in Lyre (2002, 23-27).

³ Eine andere wichtige Arbeit war Norbert Wiensers *Cybernetics, or control and communication in the animal and the machine* von 1948.

1.1.2 Kritik am Informationsbegriff in den Kognitionswissenschaften

Die Verwendung des Informationsbegriffs in den Kognitionswissenschaften ist aber in bestimmter Hinsicht problematisch. Wie wir bereits festgestellt hatten, ist der Kognitionswissenschaftler vor allem an den semantischen Eigenschaften der Informationsträger interessiert.

„Central to the idea of an information-processing mechanism is that states within the mechanism serve to carry information about objects or events external to it.“ (Bechtel 2008, 159)

Genau dieser Aspekt wird nun aber gerade vom nachrichtentechnischen Informationsbegriff überhaupt nicht erfasst. Auf diesen Punkt weist Warren Weaver, einer der beiden Verfasser von *The Mathematical Theory of Communication*, in eben diesem Werk ausdrücklich hin:

„The word information, in this theory, is used in a special sense that must not be confused with the ordinary usage. In particular, information must not be confused with *meaning*. In fact, two messages, one of which is heavily loaded with meaning and the other of which is pure nonsense, can be exactly equivalent, from present viewpoint, as regards information.“ (Weaver 1998 [1949], 8)

Für die Auffassung von Kognition als Informationsverarbeitung ist der nachrichtentechnische Informationsbegriff daher eigentlich nur von eingeschränktem Nutzen (Vgl. Hiley 2004, 199). Wortwörtlich verstanden, gibt es dementsprechend scheinbar auch keine Informationsverarbeitung durch subpersonale Prozesse, zumindest nicht, wenn wir Information als Information mit semantischem Gehalt verstehen (Vgl. Searle 2007, 113). Die unerfreuliche Gesamtsituation, die sich daraus für den Kognitionswissenschaftler ergibt, fasst Norbert Bischof sehr treffend zusammen:

"Was die Kognitionswissenschaftler aber interessiert, was insbesondere immer gemeint ist, wenn von *Verarbeitung* von Information geredet wird, ist eben gerade der Inhalt, ist die vom Signal transportierte Bedeutung, seine *semantische* Dimension. Dass ein Signal auch *Nachricht* oder *Befehl* sein kann, dass es etwas *repräsentiert*, gehörte indessen genau in die konnotative Dunstwolke, aus der Shannon den Fachausdruck "Information" wohlweislich herausgehalten hat. Damit waren die Psychologen in einer misslichen Lage. Man hatte ihnen eine Zauberformel überlassen und kleingedruckt hinzugefügt, dass sie nichts taugt. Kein Wunder, dass sie sich die Heilkraft ihrer terminologischen Placebo nicht ausreden lassen wollten und es im Zweifelsfalle vorzogen, den Dingen nicht genauer auf den Grund zu gehen. [...] Die heute so beliebte Rede vom "Informationsverarbeitungsansatz" dient lediglich als ein terminologisches Feigenblatt, das eine nach wie vor bestehende konzeptionelle Blöße bedecken soll." (Bischof 2008, 497f.)

1.1.3 Vertiefung der Kritik: Der Homunkulus-einwand

Gelegentlich wird gegen die im vorangegangenen Abschnitt aufgezeigte Schwierigkeit mit dem Informationsbegriff eingewendet, dass diese lediglich für Laien bestehe.

„In one sense there is no information in the brain at all. There are only electrical pulses and synaptic chemistry. The people using this metaphor are not misled by it. Only the lay person unfamiliar with neuroscience would draw misleading conclusions. [...] ‘Information’ in a new, mathematical sense, is well established, and should mislead no one.“ (Harré 2002, 239)

Es mag zwar zutreffen, dass Experten den umgangssprachlichen Informationsbegriff vom nachrichtentechnischen Informationsbegriff klar zu unterscheiden wissen, dies allein hilft aber auch nicht weiter, denn das eigentliche Problem besteht ja darin, dass der Kognitionsforscher selbst für seine wissenschaftlichen Erklärungen einen semantisch aufgeladenen Informationsbegriff benötigt, wie die obigen Zitate belegen. Der nachrichtentechnische Informationsbegriff kann diese Aufgabe, wie wir ebenfalls bereits gesehen haben, nicht erfüllen. Die Schwierigkeiten, die durch die Verwendung des Informationsbegriffs in den Kognitionswissenschaften entstehen, dürfen nicht unterschätzt werden. Es handelt sich hierbei nämlich nicht lediglich um kleinere konzeptionelle Ungenauigkeiten, sondern um eine Schwierigkeit, die die kognitionswissenschaftlichen Erklärungen als solche in Frage stellt. Der Vorwurf, den die Kognitionswissenschaftler in diesem Zusammenhang auf sich ziehen, ist der, einen Homunkulus-Fehlschluss zu begehen. Dieser Einwand ist keineswegs neu. Anthony Kenny hat den Ausdruck „homunculus fallacy“ geprägt:

„I shall call the reckless application of human-being predicates to insufficiently human-like objects the ‘homunculus fallacy’, since its most naïve form is tantamount to the postulation of a little man within a man to explain human experience and behaviour“ (Kenny 1991, 155).

Der Fehlschluss besteht also darin, Prädikate, die auf kognitive Leistungen von ganzen Personen angewendet werden auch auf subpersonale Prozesse zu beziehen (Vgl. Keil 2003, 78). Das eigentliche Problem beim Homunkulus-Fehlschluss ist aber nicht, dass einige intentionale Prädikate, wie „denken“, „sehen“, „verstehen“, nicht konform zu unseren normalsprachlichen Regeln verwendet werden, sondern der infinite Regress, der durch das Postulieren subpersonaler Instanzen entsteht.⁴ Wenn ich meine Denkfähigkeit dadurch erkläre, dass ich eine Instanz, einen Homunkulus, in mir habe, der denken kann, und dessen Denkfähigkeit durch einen weiteren Homunkulus im ersten Homunkulus erkläre, dann droht ein infinites Regress (Vgl. Noe 2009, 161; Keil 2003, 79; Bickhard and Terveen 1995, 173).

„The issue here, of course, is that a homunculus must itself be explained. The path of explanation implied by the homunculus idea is to reapply the same trick and suggest that another smaller homunculus might be lurking inside the first. This path leads to the specter of an infinite regress of homunculi, nested like Russian dolls, that quickly descends into absurdity.“ (Wegner 2005, 20)

⁴ Wenn ich Peter Hacker richtig verstanden habe, dann ist für ihn Ersteres wichtiger. Nach Hacker führt die Anwendung von intentionalen Prädikaten auf subpersonale Instanzen einfach zu unsinnigen Aussagen (Vgl. z.B. Bennett und Hacker 2007, 132).

Dass eine solche Erklärungsstrategie unbefriedigend ist, scheint mir offensichtlich⁵ und weniger kontrovers als der Vorwurf eines unkorrekten Gebrauchs bestimmter Ausdrücke. Der infinite Regress ist daher meines Erachtens der Kern des Homunkulus-Einwandes. Dieser Überlegung scheint auch Keil grundsätzlich zuzustimmen:

„Doch wenn ich recht sehe, kann der Homunkulismus-Kritiker die schwierige Frage nach dem originalen Anwendungsbereich des intentionalen Idioms zunächst dahingestellt sein lassen. Er muß lediglich daran festhalten, seinen Einwand im Sinne des Regreßargumentes zu formulieren. Wenn der Opponent den Anspruch erhoben hat, die kognitive Fähigkeit X reduktiv zu erklären oder zu analysieren, dann kann diese Analyse sich nicht darin erschöpfen, im Innern des fraglichen Systems ein Subsystem mit eben dieser Fähigkeit X zu identifizieren.“ (Keil 2003, 82)

Worin besteht nun aber der Zusammenhang zwischen dem Informationsbegriff und dem Homunkulus-Einwand? Von subpersonalen Instanzen wird in kognitionswissenschaftlichen Theorien ja selten behauptet, dass sie denken oder Wünsche, Überzeugungen und Absichten haben. Stattdessen ist lediglich von Informationsverarbeitung die Rede. Der dabei vorausgesetzte Informationsbegriff ist zwar nicht offen intentional, aber, so einige Kritiker, intentionalitätspräsupponierend⁶, d.h. er bezieht seinen Sinn daraus, dass in seiner Umgebung noch intentionale Phänomene unanalysiert geblieben sind (Vgl. Keil 2003, 81). Demzufolge impliziert die Auffassung von Kognition als Informationsverarbeitung zumindest indirekt, dass einige Teile unseres Zentralnervensystems miteinander kommunizieren in einem ganz ähnlichen Sinne wie auch Personen miteinander kommunizieren.

„Die Arbeitsweise subpersonaler Systeme wird nun im Idiom der Informationsverarbeitung beschrieben. Mit der Rede von Information, Beschreibung, Interpretation und Repräsentation werden einzelne kognitive Module (in den Neurowissenschaften auch physische Teile des Gehirns) als Gesprächsteilnehmer aufgefaßt, die einander Nachrichten zukommen lassen, Informationen austauschen und derart ihr Verhalten aufeinander abstimmen.“ (Keil 2003, 85)

Untergräbt also der Informationsbegriff und die Auffassung von Kognition als Informationsverarbeitung grundsätzlich die Erklärungskraft eines grossen Teils der kognitionswissenschaftlichen Theorien?

Dies ist nicht zwingend der Fall, wie wir im nächsten Abschnitt sehen werden.

⁵ Die Erklärungsstrategie des sogenannten Homunkularen Funktionalismus ist allerdings etwas subtiler. „Homunculi are *bogeymen* only if they duplicate *entire* the talents they are rung in to explain [...]. If one can get a team or committee of *relatively* ignorant, narrow-minded, blind homunculi to produce the intelligent behavior of the whole, this is progress.“ (Dennett 1978, 123) Hier ist die Erklärung nicht einfach nur verschoben, denn nun werden „höhere“ kognitive Fähigkeiten durch das Zusammenspiel von „primitiveren“ kognitiven Fähigkeiten erklärt. Dies macht den Homunkulus-Vorwurf zwar schwächer, beseitigt ihn aber letztlich nicht, denn man muss nach wie vor kognitive Fähigkeiten, wenn auch primitive, unerklärt voraussetzen.

⁶ „From the fact that a statement is not explicitly about anything mental it does not follow that none of its presuppositions make any reference to our cognitive interests, our way of regarding different contexts, or our intentional powers“ (Putnam 1992, 57). Vgl. auch Baker (1995, 193) und Dennett (1969, 23ff.).

1.1.4 Naturalistische Theorien des Gehalts als Ausweg

Unsere bisherigen Überlegungen scheinen die Kognitionswissenschaftler vor ein Dilemma zu stellen, entweder sie geben die Auffassung von Kognition als Informationsverarbeitung auf oder sie ziehen den Homunkulus-Einwand auf sich. In diesen Abschnitt möchte ich kurz darlegen, dass eine naturalistische Theorie des Gehalts, die konsistent ist, einen Ausweg aus diesem Dilemma bietet. Dafür ist es hilfreich unsere bisherigen Überlegungen zusammenzufassen. Dies waren die wesentlichen sechs Punkte:

- 1) Der Informationsbegriff ist in den Kognitionswissenschaften weit verbreitet. Kognition wird in diesen Wissenschaften überwiegend als Informationsverarbeitung aufgefasst.
- 2) Der Informationsbegriff in den Kognitionswissenschaften ist semantisch aufgeladen. Die Informationsträger, die bei der als Informationsverarbeitung verstandenen Kognition eine Rolle spielen sollen, haben einen semantischen Gehalt.
- 3) Das Aufkommen des „Informationsparadigmas“ (Bischof 2009, 493) in den Kognitionswissenschaften hat seinen Ursprung in der Entwicklung des nachrichtentechnischen Informationsbegriffs durch Shannon und Weaver. Der nachrichtentechnische Informationsbegriff ist objektiv und naturwissenschaftlich respektabel, d.h. frei von Metaphysikverdacht.
- 4) Der nachrichtentechnische Informationsbegriff hat keine semantische Dimension, er bezieht sich nur auf die syntaktischen Eigenschaften einer Zeichenkette. Der nachrichtentechnische Informationsbegriff ist daher für die Auffassung von Kognition als Informationsverarbeitung völlig unzureichend (aus 2).
- 5) Wenn in den kognitionswissenschaftlichen Theorien nicht der nachrichtentechnische Informationsbegriff gemeint sein kann (aus 4), dann liegt es nahe „Information“ in diesen Theorien nicht als technischen Ausdruck, sondern umgangssprachlich zu verstehen.
- 6) Ein umgangssprachliches Verständnis des Informationsbegriffs in den kognitionswissenschaftlichen Theorien stellt deren Erklärungskraft aufgrund des Homunkulus-Einwandes in Frage.

In Punkt 5, der in unseren Überlegungen nur indirekt eine Rolle gespielt hat, wird deutlich, worin ein möglicher Ausweg bestehen könnte. Es ist ja nicht so, dass der Informationsbegriff in den Kognitionswissenschaften nur auf zweierlei Weise verstanden werden kann, entweder nachrichtentechnisch oder umgangssprachlich. Der Ausweg besteht also darin ein drittes Verständnis anzubieten, indem man einen Informationsbegriff entwickelt, der eine semantische Dimension hat und dennoch nicht den Homunkulus-Einwand heraufbeschwört. Mit anderen Worten: Benötigt wird eine Theorie, die angibt, was Informationsträger sind und welchen semantischen Gehalt diese Träger haben, ohne dabei kognitive Fähigkeiten, wie Verstehen, Erkennen etc., vorauszusetzen. Naturalistische Theorien des Gehalts sind Theorien, die beanspruchen diese Anforderung zu erfüllen (Vgl. Rupert 2009, 236f; Keil 2003, 92)

„Thus, a third strategy for handling the regress worry has become quite popular. This strategy has been to offer an account of representation that drops the sophisticated homunculus, but nonetheless retains enough of the relevant relations, conditions and properties of representation so that, in the end, we still have something that is recognizably a representational system. In other words, the strategy is to show that we

can have something that functions as a representation in a physical system, even if there is no sophisticated built-in learner or inference-maker that it serves as a representation for. [...] With this method, the way to understand mental representation involves coming up with an account of representation that requires nothing playing the role of an inner mind that uses the representation.“ (Ramsey 2007, 192f)⁷

Damit ist der Kontext, in dem naturalistische Theorien des Gehalts meines Erachtens eine bedeutende Rolle spielen, umrissen.

In den nächsten Abschnitten soll näher erläutert werden, was naturalistische Theorien des Gehalts sind, indem folgende Fragen beantwortet werden:

- Was ist das Ziel naturalistischer Theorie des Gehalts?
- Sind naturalistische Theorien des Gehalts überhaupt „Theorien“ im wissenschaftstheoretischen Sinn?
- Welcher Anspruch ist mit dem Zusatz „naturalistisch“ verbunden?
- Was ist mit Gehalt gemeint?

⁷ In den Kognitionswissenschaften werden die Träger der Informationen oft als Symbole oder Repräsentationen bezeichnet. In diesen Sinne spricht Rupert bezogen auf die hier verhandelte Problematik auch vom “Symbol-Grounding Problem” (Rupert 2009, 236).

1.2 Naturalistische Theorien des Gehalts

1.2.1 Das Ziel naturalistischer Theorien des Gehalts

Stephen Stich hat zu Recht moniert, dass in den Debatten um naturalistische Theorien des Gehalts oft nicht explizit formuliert wird, was eigentlich das Ziel solcher Theorien ist und ab wann das Ziel als erreicht gilt.

„While there is no shortage of debate about the merits and demerits of various accounts of mental content, there has been remarkably little discussion about what a theory of mental representation is supposed to do: What question (or questions) is a theory of mental representation supposed to answer? And what would count as getting the answer right?“ (Stich 1992, 243f)

In diesem Abschnitt soll deshalb die erste Frage, die nach dem Ziel naturalistischer Theorien des Gehalts, beantwortet werden. Der zweiten Frage, ab wann das Ziel als erreicht gilt, wenden wir uns später zu.

Gehen wir zunächst allgemein von Theorien des Gehalts aus. Deren Ziel bestimmt Neander wie folgt:

„In general, a theory of content tries to say why a mental representation counts as representing what it represents.“ (Neander 2012)

Etwas genauer formuliert, kann man sagen, dass es das Ziel einer Theorie des Gehalts ist, in allgemeiner Weise anzugeben, was den Gehalt eines Gehaltsträgers bestimmt, so dass mit Hilfe der Theorie jedem beliebigen Gehaltsträger ein bestimmter Gehalt zugeordnet werden kann. Eine solche Theorie des Gehalts könnte also zum Beispiel besagen, dass die Ursache eines Gehaltsträgers zugleich sein Gehalt ist. Auf diese Weise wäre der Gehalt eines jeden Gehaltsträgers festgelegt.

Das Ziel naturalistischer Theorien des Gehalts muss aber meines Erachtens noch umfassender verstanden werden. In diesen Theorien geht es oder sollte es zumindest zudem darum gehen, die Frage zu beantworten, was überhaupt Gehaltsträger sind. Beantwortet man diese Frage nicht, dann wird man unweigerlich und zu Recht Gehaltsträger einfach mit *unseren* Gedanken oder Ausdrücken *unserer* Sprache assoziieren. Ist dies erst einmal geschehen, lässt sich der Homunkulus-Einwand nicht mehr vermeiden. Es ist also für naturalistische Theorien des Gehalts sehr wichtig, dass sie die Frage, was überhaupt ein Gehaltsträger ist, beantworten, zumindest wenn sie einen Ausweg aus dem oben skizzierten Dilemma bieten wollen. Diesen Zusammenhang sieht auch Keil, der explizit darauf hinweist, dass die Triftigkeit des Homunkulus-Einwandes von der Beantwortung der Frage abhängt, was überhaupt mentale Repräsentationen sind.

„Mündet die Annahme „mentaler Repräsentationen“ zwangsläufig in den Homunkulismus? Die Antwort hängt unter anderem davon ab, was mentale Repräsentationen überhaupt sind.“ (Keil 2003, 88)

Nach dem hier vertretenen Verständnis von naturalistischen Theorien des Gehalts verfolgen diese Theorien also zwei Ziele:

- Ziel 1: eine allgemeine Bestimmung davon, was ein Gehaltsträger ist
- Ziel 2: eine allgemeine Angabe, darüber, wodurch der Gehalt von Gehaltsträgern festgelegt ist

1.2.2 Theorieverständnis

In diesem Abschnitt geht es darum zu zeigen, dass naturalistische Theorien des Gehalts als „Theorien“ im wissenschaftstheoretischen Sinn verstanden werden können⁸. Hintergrund dieses Abschnittes ist die unter Philosophen weit verbreitete Auffassung, dass Philosophen keine Theorien im wissenschaftstheoretischen Sinn aufstellen⁹.

Innerhalb der Wissenschaftstheorie selbst gibt es allerdings verschiedene Auffassungen darüber, was Theorien sind (Vgl. Thiel 1996, 267-68; Giere 2000, 215-220). Ich beschränke mich hier darauf darzulegen, dass naturalistische Theorien des Gehalts zumindest dem Theorieverständnis der Wissenschaftstheoretiker Martin Mahner und Mario Bunge entsprechen¹⁰. Nach Mahner und Bunge ist eine Theorie „ein *hypothetiko-deduktives System*, ein System von Hypothesen, die durch die logische Relation der Deduktion miteinander in Beziehung stehen.“ (Mahner und Bunge 2000, 81) Ein solches hypothetiko-deduktives System wird bereits durch eine einzelne All-Aussage und deren Folgesätze gebildet.

„Dahingegen ist die Hypothese „Alle x sind Fs, wobei x zu S gehört“, zusammen mit all ihren unmittelbaren Folgesätzen, ein (bescheidenes) hypothetiko-deduktives System. Es bildet jedoch keine wissenschaftliche Theorie, weil es keine Gesetzesaussagen enthält. Da die Hypothese „Alle x sind Fs, wobei x zu S gehört“ die Grundaussage dieses hypothetiko-deduktiven Systems ist, aus der alle anderen Aussagen folgen, wird sie als *Axiom* oder *Postulat* bezeichnet.“ (Mahner und Bunge 2000, 83)¹¹

Wir hatten bereits dargelegt, welche Ziele naturalistische Theorien des Gehalts verfolgen. Insbesondere das Ziel, in allgemeiner Weise anzugeben, wodurch der Gehalt eines jeden Gehaltsträgers festgelegt ist, erfordert die Formulierung mindestens einer All-Aussage. Diese All-Aussage könnte z.B. behaupten, dass die Ursache eines jeden Gehaltsträgers zugleich dessen Gehalt ausmacht. Diese Aussage wäre dann ein Postulat und bildet zusammen mit Einzel-Aussagen, wie „x ist Ursache y“ und „y ist ein Gehaltsträger“, ein hypothetiko-deduktives System. Dieses hypothetiko-deduktive System kann dann nach Mahner und Bunge als rudimentäre Theorie verstanden werden.

Dieses Theorieverständnis rückt Theorien in die Nähe von einzelnen All-Aussagen wie „Alle Wale säugen ihre Jungen“ (Vgl. Nagel 1961, 79-81)¹² und ist in der Wissenschaftstheorie durchaus gängig. Für Hempel und Oppenheim sind sogar einzelne generalisierte Existenz-Aussagen „Theorien“ (Vgl. Salmon 2006 [1989], 194).

Im Allgemeinen erschöpfen sich wissenschaftliche Theorien natürlich nicht in einem einzelnen Postulat, sondern sie bestehen aus mehreren Postulaten. Von den Postulaten einer

⁸ Thiel unterscheidet vier Verwendungsweisen des Wortes „Theorie“: 1) eine außerwissenschaftliche, 2) eine philosophische, 3) eine einzelwissenschaftliche und 4) eine wissenschaftstheoretische (Vgl. Thiel 1996, 260).

⁹ Diese Auffassung wurde vor allem von Wittgenstein vertreten. „In his later writings Wittgenstein holds, as he did in the *Tractatus*, that philosophers do not—or should not—supply a theory, neither do they provide explanations. ‘Philosophy just puts everything before us, and neither explains nor deduces anything.—Since everything lies open to view there is nothing to explain’ (PI 126)” (Biletzki and Matar 2014).

¹⁰ Ich beziehe mich hierbei vor allem auf den Abschnitt 3.5 *Theorie und Modell* in Mahner und Bunge (2000, 81-99).

¹¹ Darüber, was genau Gesetze bzw. Gesetzesaussagen sind, besteht ebenfalls Uneinigkeit (Vgl. Carroll 2012).

¹² Die Aussage über Wale bezeichnet Nagel auch als experimentelles Gesetz und unterscheidet diese von theoretischen Gesetzen oder einfachen Theorien. Die Unterscheidung ist aber keine strikte, da diese wiederum auf der unklaren Unterscheidung zwischen „beobachtbar“ und „nicht beobachtbar“ beruht. „In consequence, the distinction [between experimental laws and theories, A.W.] suffers from all the notorious unclarities that attach to the word “observable”.” (Nagel 1961, 81).

Theorie wird dann gefordert, dass sie voneinander logisch unabhängig sind und dass sie miteinander kompatibel sind (Vgl. Salmon 2006 [1989], 177)¹³. Vor allem die Widerspruchsfreiheit gilt als *conditio sine qua non* für die Bewertung einer Theorie (Vgl. Thiel 1996, 265).

Neben dieser grundlegenden Forderung an die Postulate einer Theorie, spielt für die Bewertung einer Theorie auch die Möglichkeit ihrer prinzipiellen empirischen Überprüfbarkeit eine wichtige Rolle. Nach Mahner und Bunge müssen für die Überprüfung einer Theorie an der Realität sogenannte Indikatorhypothesen für die theoretischen Größen der Theorie aufgestellt werden. Sie erläutern dies an einem einfachen Beispiel. Angenommen unsere Theorie besagt, dass Stress hinreichend für Kopfschmerzen ist (Substantielle Hypothese). Die Größe „Kopfschmerz“ kann man direkt fühlen, die Größe „Stress“ hingegen nicht. „Stress“ betrachten Mahner und Bunge, daher als eine theoretische Größe. Ein hoher Stresslevel wird jedoch durch einen hohen Kortikosteronergehalt im Blut angezeigt (Indikatorhypothese¹⁴). Aus der substantiellen und der Indikatorhypothese folgt nun, dass ein hoher Kortikosteronergehalt hinreichend für Kopfschmerzen ist. Dieser Zusammenhang ist nun durch Beobachtung überprüfbar (Mahner und Bunge 2000, 91f)¹⁵.

Indikatorhypothesen lassen sich natürlich auch für naturalistische Theorien des Gehalts aufstellen. Eine solche Indikatorhypothese könnte z.B. lauten, dass der Gehalt eines Gehaltsträgers durch bestimmte Verhaltensäußerungen angezeigt wird¹⁶. Naturalistische Theorien des Gehalts können daher prinzipiell auch an der Realität scheitern. Eine solche empirische Überprüfung bedarf aber vertrauenswürdiger Indikatorhypothesen. Mit anderen Worten: Solange die Indikatorhypothese ebenso fragwürdig ist, wie die substantielle Hypothese, bleibt die Überprüfung uneindeutig.

Insgesamt können wir also festhalten, dass keinen Grund gibt, warum naturalistische Theorien des Gehalts nicht als Theorien im wissenschaftstheoretischen Sinne verstanden werden können. Zudem können naturalistische Theorien des Gehalts wie andere Theorien auch aus mindestens zwei Gründen scheitern:

- 1) die Postulate der Theorie sind nicht miteinander kompatibel
- 2) die Theorie hält einer Überprüfung an der Realität nicht stand

Der erste der beiden Gründe ist dabei insofern vorrangig als es konsistente Theorien gibt, die einer Überprüfung an der Realität nicht standhalten, aber es gibt keine inkonsistenten Theorien, die einer Überprüfung an der Realität standhalten. Anders ausgedrückt: Inkonsistente Theorie sind in jeden Fall gescheitert, eine Überprüfung, ob sie auch der Realität standhalten, lohnt sich gar nicht erst. Aus diesem Grund ist bei einer Theorie zunächst zu prüfen, ob sie konsistent ist.

¹³ Salmon referiert an dieser Stelle Watkins. Watkins hält die genannten Forderungen zwar für korrekt aber nicht für ausreichend, denn zwei Aussagen, wie „Alle Raben sind schwarz“ und „Alle Krokodille sind dickhäutig“, erfüllen die Forderungen, bilden aber dennoch keine Theorie (Vgl. Watkins 1984, 204). Auf Watkins Versuch dieses Problem zu lösen, kann hier nicht eingegangen werden.

¹⁴ Indikatorhypothesen sind nicht Teil der Theorie (Vgl. Mahner und Bunge 2000, 92-94).

¹⁵ Nach Mahner und Bunge sind noch weitere Schritte nötig um eine Theorie empirisch zu überprüfen (Vgl. Mahner und Bunge 2000, 92-94). Auf diese Schritte soll hier nicht eingegangen werden.

¹⁶ Diese Hypothese ist sicherlich nicht besonders exakt formuliert. An dieser Stelle geht es aber auch nur darum, dass es plausibel ist, dass solche Indikatorhypothesen auch in Bezug auf naturalistische Theorien des Gehalts formuliert werden können.

1.2.3 Naturalismusverständnis

In diesem Abschnitt gilt es zu klären, welcher Anspruch in dieser Arbeit mit dem Zusatz „naturalistisch“ für eine Theorie des Gehalts verbunden werden soll. Es geht nicht darum, ausführlich zu beleuchten, was alles unter „Naturalismus“ verstanden wird.

Obwohl sich viele Philosophen zum Naturalismus bekennen, bleibt der Inhalt der Bekenntnisse oft vage (Vgl. Godfrey-Smith 2003, 150). Mit Naturalismus ist häufig ganz allgemein ein philosophischer Ansatz gemeint, der die Verbindung zwischen Philosophie und Naturwissenschaft betont (Vgl. Godfrey-Smith 2003, 150). Dementsprechend konturlos bleibt dann auch die so verstandene naturalistische Position.

"The view I favor is Liberal Naturalism. I view naturalism as a methodological requirement to place human beings in the world without making special, ad hoc assumptions that are discontinuous with everything else we have good reason to believe about nature." (Rosenberg 2004, 8)

Darüber hinaus wird häufig zwischen einem ontologischen und einem methodologischen Naturalismus unterschieden. Der ontologische Naturalismus besagt, dass alles, was es gibt, Teil der Natur ist, während der methodologische Naturalismus behauptet, dass alle Phänomene durch naturwissenschaftliche Methoden erfasst und erklärt werden können (Vgl. Keil 2000, 188). Teilweise wird der Naturalismus auch als Konjunktion der beiden Formen bzw. Thesen verstanden:

"*naturalism*, the twofold view that (1) everything is composed of natural entities – those studied in the sciences (on some versions, the natural sciences) – whose properties determine all the properties of things, persons included (abstracta like possibilities and mathematical objects, if they exist, being constructed of such abstract entities as the sciences allow); and (2) acceptable methods of justification and explanation are continuous, in some sense, with those in science. Clause (1) is metaphysical or ontological, clause (2) methodological and/or epistemological." (Post 1999, 596)¹⁷

Beide Formen des Naturalismus kranken an einer Unterbestimmtheit der entscheidenden Begriffe. Der ontologische Naturalismus kann nicht genau angeben, was *das Natürliche* ist (Vgl. Strawson 2010, 2; Keil 2000, 200) und der methodologische Naturalismus kann nicht genau angeben, was *die Methoden* der Naturwissenschaften ausmacht (Vgl. Keil 2000, 200).

Um zu klären, welcher Anspruch mit dem Zusatz „naturalistisch“ für eine Theorie des Gehalts verbunden werden soll, werde ich daher die genannten Formen des Naturalismus unberücksichtigt lassen und mich stattdessen an einem Vorschlag von Geert Keil orientieren.

Nach Keil ist eine naturalistische Theorie dadurch gekennzeichnet, dass sie im Explanans auf das intentionale Idiom und dessen Präsuppositionen verzichtet.

„Das Kriterium hat die Form einer Ausschlußbedingung. Naturalistisch akzeptierbar seien nur solche Erklärungen oder Theorien, die nicht auf das intentionale Idiom der Alltagssprache zurückgreifen.“ (Keil 2000, 190)

Eine ganz ähnliche Ausschlußbedingung ergab sich auch in Bezug auf die Vermeidung des Homunkulus-Einwandes. Um dort den infiniten Regress zu vermeiden, war es erforderlich die Erklärung einer kognitiven Fähigkeit eines Organismus' nicht auf irgendwelche denkenden Instanzen innerhalb dieses Organismus' zu stützen. Wenn also eine Theorie des Gehalts einen möglichen Ausweg zur Vermeidung des Homunkulus-Einwandes sein soll,

¹⁷ Vgl. auch Rakova (2006, 126).

dann darf sie dem Einwand trivialerweise nicht selbst erliegen. Der Zusatz „naturalistisch“ in „naturalistische Theorie des Gehalts“ soll in dieser Arbeit dementsprechend den Anspruch beinhalten, dass für das Erreichen der zwei Ziele einer naturalistischen Theorie des Gehalts keine denkenden Instanzen vorausgesetzt werden dürfen. D.h., weder die Bestimmung, was ein Gehaltsträger ist, noch die Angabe, was den Gehalt dieser Gehaltsträger bestimmt, darf irgendwelche denkenden Instanzen in Anspruch nehmen. Dieser Anspruch ist in dem Kontext, in dem wir die naturalistischen Theorien des Gehalts gestellt haben, folglich nichts anderes als der Anspruch, den infiniten Regress zu vermeiden. Die mit diesem Anspruch verbundene Forderung scheint nicht nur sinnvoll, sondern fast schon selbstverständlich.

„Cognitive science hopes to explain what it is to understand an utterance by appeal to the processing of representations and indicator signals, so the notion of content applied to representations and indicator signals had better not presuppose understanding anything.“ (Cummins and Roth 2010, 177)

1.2.4 Gehaltsverständnis

In diesem Abschnitt soll geklärt werden, was wir in dieser Arbeit unter dem Gehalt eines Gehaltsträgers verstehen wollen.

Hierfür können wir uns zunächst am sprachlichen Zeichen als einem paradigmatischen Gehaltsträger orientieren. Damit rücken wir die naturalistische Theorie des Gehalts in die Nähe von Bedeutungstheorien sprachlicher Ausdrücke. Fragen nach dem Gehalt eines Gehaltsträgers entsprechen dort Fragen nach der Bedeutung eines sprachlichen Ausdrucks.

Seit Freges Aufsatz „Über Sinn und Bedeutung“ von 1892 unterscheidet man hinsichtlich der Bedeutung eines Ausdrucks zwei Aspekte, seinen Sinn und seine Bedeutung¹⁸ (Referenz). In abweichender Terminologie finden wir diese Unterscheidung auch bei Carnap, Morris und Quine. Carnap unterscheidet zwischen Intension und Extension, Morris zwischen Designation und Denotation und Quine zwischen meaning und reference (Vgl. Lyre 2002, 180). Diese Theorie der doppelten Bedeutung lässt sich auch durch das sogenannte semiotische Dreieck veranschaulichen (Vgl. Lorenz 1984, 769).

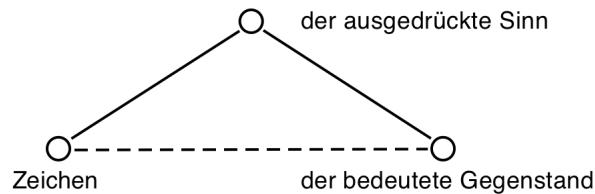


Abbildung 1: Semiotisches Dreieck (In Anlehnung an Lorenz 1984, 769)

So ist nach Frege die Bedeutung (Referenz) eines Namens, wie „der Morgenstern“, ein Gegenstand, in diesem Fall der Planet Venus. Der Sinn des Namens ist hingegen nach Frege „die Art des Gegebenseins“ (Frege 1990 [1892], 143f [26]) des betreffenden Gegenstandes. Der Sinn eines Ausdrucks ist also so etwas wie eine „identifizierende Bedingung“ (Stepanians 2001, 151). Im Fall des Ausdrucks „der Morgenstern“ könnte diese Bedingung in etwa lauten „Stern, der am Morgen (dort und dort am Firmament) zu sehen ist“. D.h., indem der Hörer den Sinn des Ausdrucks erfasst, ist er in der Lage den Gegenstand zu identifizieren, der die Bedeutung (Referenz) des Ausdrucks ist. Nach Frege bestimmt also der Sinn die Bedeutung eines Ausdrucks. Dies gilt darüber hinaus nicht nur für Namen und Kennzeichnungen, sondern auch für Begriffsworte (Prädikate) und Sätze (siehe Abbildung 2).

¹⁸ Frege verwendet den Ausdruck „Bedeutung“ etwas eigentümlich, heute spricht man eher von Referenz.

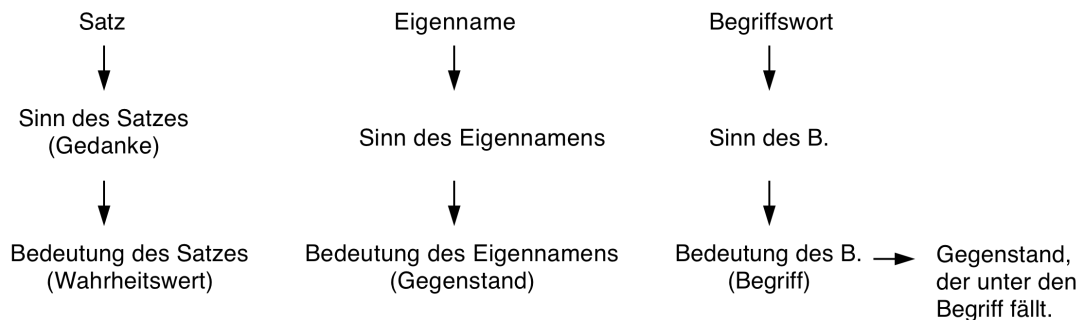


Abbildung 2: Sinn und Bedeutung sprachlicher Ausdrücke nach Frege (In Anlehnung an Frege 1976 [1891], 96)

Freges Bedeutungstheorie besteht also allgemein gesprochen aus zwei theoretischen Ebenen bzw. aus zwei Theorien.

„We can think of the theory of Meaning [Bedeutung/Referenz, A.W.], very generally, as concerned with relations between expressions of the language and entities in the world. Now, Frege came to see the need for another level of description and theory which is concerned, we might with equal generality say, with relations between expressions of the language and the understanding that competent speakers have of them“ (Evans 1982, 13)

In dieser Weise wird innerhalb der Semantik oft zwischen einer Theorie der Referenz und einer Theorie der Bedeutung unterschieden.

„The broad field of semantics is divided into three subjects [...], the first is syntax, the second is semantics, and third is pragmatics. The narrower field of semantics is sometimes further divided [...] into two subjects: the theory of reference (denotation, extension) and theory of meaning (connotation, intension).“ (Kalish 1972, 348)

Welcher der beiden Theorie dabei das Primat zukommt, ist umstritten (Vgl. Lorenz 1984, 771).

Kommen wir nun auf die Frage zurück, was wir in dieser Arbeit unter dem Gehalt eines Gehaltsträgers verstehen wollen. Der Gehalt eines sprachlichen Ausdrucks hat, wie soeben dargelegt, zwei Aspekte:

- 1) seine Referenz
- 2) seinen Sinn

In dieser Arbeit soll nun unter dem Gehalt eines Gehaltsträgers in einer naturalistischen Theorie des Gehalts *nur der erste Aspekt, die Referenz*, verstanden werden. Dies bedeutet insbesondere, dass es bei dem zweiten Ziel einer naturalistischen Theorie des Gehalts, nämlich allgemein anzugeben, was den Gehalt eines Gehaltsträgers bestimmt, darum geht anzugeben, was allgemein die Referenz eines Gehaltsträgers bestimmt¹⁹.

Diese Einschränkung ist durch zwei Gründe motiviert. Der erste Grund ist rein praktischer Natur. Aufgrund des Umfangs dieser Untersuchung ist einfach eine Einengung der Fragestellung notwendig.

¹⁹ Die Theorie übernimmt dabei also eine analoge Rolle wie Freges Sinn, denn nach Frege bestimmt der Sinn die Bedeutung. Da Freges Sinnbegriff aber mindestens epistemisch aufgeladen ist (Vgl. Glock 2000), kann er für eine naturalistische Theorie nicht in Frage kommen.

Der entscheidende Grund ist aber der zweite und dieser ist inhaltlicher Natur. Wie wir bereits gesehen haben, hängt der Sinn eines Ausdrucks und ganz allgemein die Sinnebene sehr eng mit dem verstehenden Erfassen eines Ausdrucks durch ein geistbegabtes Wesen zusammen (Vgl. auch Fussnote 20). Frege führt die Unterscheidung zwischen Sinn und Bedeutung ja überhaupt erst ein, um so den unterschiedlichen Erkenntniswert von Identitätsaussagen, wie „a=a“ und „a=b“, erklären zu können (Vgl. Frege 1990 [1892], 143 [26]).

Wenn also die Auffassung von Gehalt als Sinn im vornherein an ein verstehendes Wesen gebunden ist bzw. ein solches voraussetzt, dann darf eine naturalistische Theorie des Gehalts, die es ermöglichen soll Kognition als Informationsverarbeitung zu verstehen ohne dem Homunkulus-Einwand zu erliegen, Gehalt nicht als Sinn begreifen. Wie wir gesehen haben, beruhte der Homunkulus-Einwand ja genau darauf Information, Gehalt und Gehaltsträger in einer Weise zu verstehen, die von unserer menschlichen Kommunikation mittels sprachlicher und anderer Zeichen ausgeht. Es scheint dementsprechend aussichtslos mittels einer naturalistischen Theorie des Gehalts dem Homunkulus-Einwand zu entkommen wollen, wenn diese Gehalt auch als Sinn zu versteht. Die Einschränkung ist also sinnvoll (wenn nicht gar notwendig), insofern in dieser Arbeit naturalistische Theorien des Gehalts hinsichtlich ihrer Fähigkeit betrachtet werden eine ganz bestimmte Aufgabe zu erfüllen.

Ebenso kann sich eine naturalistische Theorie des Gehalts auch mit ihrem ersten Ziel, nämlich der Bestimmung dessen, was ein Gehaltsträger überhaupt ist, nicht einfach an sprachlichen Zeichen orientieren, zumindest nicht, wenn sie dem Homunkulus-Einwand entgehen will. Vielmehr muss eine naturalistische Theorie des Gehalts zwingend einige Aspekte der Bedeutungstheorien aufgeben, denn es geht ja gerade darum, ein Verständnis von Information, Gehalt und Gehaltsträger zu entwickeln, das von unserem bisherigen, an der menschlichen Kommunikation orientierten, verschieden ist.

Dass in einer naturalistischen Theorie des Gehalts aber immer noch von Gehalt und nicht von etwas gänzlich Neuem die Rede ist, wird – so jedenfalls die Hoffnung – dadurch gewährleistet, dass zumindest der erste Aspekt, die Referenz, in der modifizierten Gehaltsauffassung enthalten bleibt. Damit bleibt die modifizierte Auffassung von Information, Gehalt und Gehaltsträger aber andererseits auch nicht, wie die informationstheoretische Auffassung, auf die syntaktische Dimension beschränkt, sondern hat durchaus eine semantische Dimension (eben auch im Sinne der Bedeutungstheorien sprachlicher Ausdrücke). Soweit zur Begründung der Einschränkung.

Ich möchte jetzt noch kurz auf zwei mögliche Konsequenzen dieser Einschränkung für die Anwendung einer naturalistischen Theorie des Gehalts auf sprachliche Ausdrücke eingehen. Einerseits könnte man sagen, dass eine naturalistische Theorie des Gehalts für sprachliche Ausdrücke ungeeignet oder zumindest unvollständig ist, weil sie einen wichtigen Aspekt des Gehaltes bzw. der Bedeutung, nämlich den Sinn, unberücksichtigt lässt. Andererseits könnte man auch dafür argumentieren, dass naturalistische Theorien des Gehalts für sprachliche Ausdrücke durchaus geeignet sind, weil der Bedeutungsaspekt des Sinns völlig unwesentlich und vernachlässigbar oder gar eine Fehlkonzeption ist.

Die zweite Argumentationsstrategie verfolgt Millikan. Wir werden daher auf diese Option unter Abschnitt 1.4.3 *Millikans Anspruch* noch einmal zurückkommen.

Ich möchte aber ausdrücklich betonen, dass die Beantwortung der Frage, ob eine naturalistische Theorie des Gehalts eine Alternative zu Bedeutungstheorien sprachlicher Ausdrücke darstellt, nicht Teil dieser Arbeit ist.

Diesen Abschnitt abschliessend mag noch ein Hinweis zur Terminologie in dieser Arbeit angebracht sein. Einigen Lesern mag es merkwürdig erscheinen, eine um den Sinneaspekt reduzierte Theorie immer noch als Theorie des Gehalts zu bezeichnen. Der Gehalt (content) ist doch nahezu identisch mit dem Sinn (sense) (Vgl. z.B. Crane 2001, 28-33). Warum ist hier also nicht die Rede von einer Theorie der Referenz anstatt irreführend von einer Theorie des

Gehalts zu reden? Der Grund ist einfach folgender: Millikans Theorie, insbesondere diejenigen Aspekte, die uns hier interessieren, werden in der Literatur unter dem Begriff „theory of content“ diskutiert (Vgl. z.B. Neander 2012, Fodor 1990). Dementsprechend finden wir in den deutschen Übersetzungen von Millikan und der deutschsprachigen Sekundärliteratur „Theorie des Gehalts“ bzw. „Gehaltstheorien“ (Vgl. z.B. Millikan 2008, Detel 2001a). Ich folge hier dem in der Diskussion üblichen Sprachgebrauch. Ein Wechsel der Terminologie würde auf viele Leser noch verwirrender wirken.

Grundlegend unterscheide ich hier nur zwischen „Gehaltsträger“ und „Gehalt“, wobei „Gehalt“ im zuvor dargelegten reduzierten Sinn zu verstehen ist. In der Arbeit gebrauche ich weitere Begriffe, wie „Signal“, „Repräsentation“, „Information“, die sich jeweils einem der beiden Begriffe unterordnen lassen. Ich habe mich gegen eine einheitliche Terminologie entschieden, weil sich einige Begriffe in bestimmten Bereich durchgesetzt haben und eine Vereinheitlichung hier seltsam anmuten würde.²⁰ Aus der folgenden Tabelle kann aber eine Zuordnung der einzelnen Begriffe zu den beiden grundlegenden Begriffen „Gehaltsträger“ und „Gehalt“ entnommen werden.

Gehaltsträger	Gehalt
Informationsträger, Signal, Repräsentation ²¹ , Zeichen, Icon	Information, Bedeutung

²⁰ Andernfalls müsste ich von „Tiergehaltsträger“ anstatt von „Tiersignalen“ sprechen etc.

²¹ Insbesondere der Begriff der Repräsentation wird in Literatur doppeldeutig gebraucht, teilweise bezeichnet er den Gehalt und teilweise den Träger des Gehalts (Vgl. Glock 2014, 527-531).

1.2.5 Bedingung: Missrepräsentation muss möglich sein

In diesem Abschnitt geht es darum eine wichtige Grundannahme naturalistischer Theorien des Gehalts einzuführen und zu rechtfertigen. Die Grundannahme lautet: Es ist möglich, dass der Gehalt eines Gehaltsträgers nicht mit der Realität übereinstimmt. D.h. der Gehalt eines Gehaltsträgers muss falsch bzw. inkorrekt sein können. Kurz: Missrepräsentation muss möglich sein.

Diese Grundannahme wird als *condition sine qua non* sowohl von Naturalisten als auch von ihren Gegnern nahezu ausnahmslos geteilt (Vgl. z.B. Ramsey 2007, 130; Rakova 2006, 118; Perlman 2000, 1, 31).

„Much attention is paid to the possibility of misrepresentation. This is because the distinction between correct and incorrect representation is often regarded as a central normative distinction and because a capacity to misrepresent is often thought to be essential for representing: no possibility of misrepresentation, no representing.“
(Neander 2012)

Mark Perlmann, einer der wenigen, der bereit ist diese Annahme aufzugeben, bezeichnet diese deshalb auch als „dogma of meaning“ (Perlman 2000, 1).

Die Annahme, dass Missrepräsentation möglich sein muss, wird häufig durch unsere eigenen Erfahrungen gerechtfertigt. Wir wissen einfach, dass Missrepräsentationen möglich sind, weil wir selbst erlebt haben, dass uns unsere Wahrnehmung getäuscht hat oder dass unsere eigenen Überzeugungen falsch waren (Vgl. z.B. Matthen 2006, 149; Neander 2012; Rakova 2006, 118). Eine weitere Rechtfertigungsmöglichkeit ist begrifflicher Natur. Sie besteht darin zu behaupten, dass es zum Begriff eines Gehalts- bzw. Informationsträgers oder einer Repräsentation gehört, dass deren Gehalt bzw. Information auch falsch bzw. inkorrekt sein kann (Vgl. z.B. Ramsey 2007, 32).

Ich halte beide Rechtfertigungen an sich für überzeugend. Sie sind allerdings in dem Kontext, in den wir die naturalistischen Theorien des Gehalts gestellt haben, weniger zwingend und bedürfen zusätzlicher Bemerkungen.

Die begriffliche Rechtfertigung geht von unserem alltagssprachlichen, vorthoretischen Verständnis der Begriffe „Information“, „Gehalt“ und „Repräsentation“ aus. Dieses Verständnis orientiert sich im Wesentlichen an der menschlichen Kommunikation mittels sprachlicher Zeichen. Im vorangegangenen Abschnitt hatten wir aber bereits gesehen, dass die naturalistischen Theorien des Gehalts von diesem Verständnis abweichen müssen, wenn sie dazu beitragen sollen, dass die Auffassung von Kognition als Informationsverarbeitung dem Homunkulus-Einwand entgeht. Aus diesem Grund ist die begriffliche Rechtfertigung in unserem Kontext nicht unmittelbar zwingend, denn wir könnten neben der Sinnebene, die ein geistbegabtes Wesen voraussetzt, auch die Möglichkeit zur Missrepräsentation als konstitutiv für den Begriff der Repräsentation (des Gehalts- bzw. Informationsträger) aufgeben. Letzteres scheint mir aber aus zwei Gründen nicht sinnvoll. Erstens, setzt die Annahme, dass Missrepräsentation möglich sein muss, zumindest nicht unmittelbar das Verstehen des Gehalts durch ein geistbegabtes Wesen voraus²² und ist daher auch nicht unbedingt aufzugeben. Im Gegenteil sie ist zur Wahrung der Kontinuität mit unserem

²² Vgl. aber Peter Slezaks "The World Gone Wrong? Images, Illusions, Mistakes and Misrepresentations." (Slezak 2004). „He [Peter Slezak, A.W.] proposes the intriguing and challenging idea that talk of misrepresentation commits the homunculus fallacy.“ (Clapin 2004, xxii).

vortheoretischen Verständnis der Begriffe „Information“, „Gehalt“ und „Repräsentation“ eher beizubehalten. Damit hängt auch der zweite Grund zusammen. Die Annahme, dass Missrepräsentation möglich sein muss, scheint mir zumindest intuitiv grundlegender mit unserem vortheoretischen Verständnis der Begriffe „Information“, „Gehalt“ und „Repräsentation“ verbunden zu sein als die Sinnebene der Bedeutung. So hatte auch Frege zunächst eine Bedeutungstheorie entwickelt, die ohne die Unterscheidung zwischen Sinn und Bedeutung auskam, bei der aber die Wahrheit und damit auch die Falschheit von Sätzen eine zentrale Rolle spielte (Vgl. Evans 1982, 8-14).

Die Rechtfertigung der Annahme, dass Missrepräsentation möglich sein muss, durch unsere eigenen Erfahrungen kann ebenfalls nicht ohne Weiteres in den hier betrachteten Kontext übertragen werden. Diese Rechtfertigung geht davon aus, dass Repräsentationen mentale Repräsentationen sind und versteht diese wiederum als Sinneserfahrungen oder Gedanken, die man phänomenologisch als Bewusstseinsepisoden bezeichnen würde²³. Von einer solchen Gleichsetzung kann in dem hier abgesteckten Rahmen natürlich nicht ausgegangen werden. Gehaltsträger bzw. Repräsentationen werden bei der Auffassung von Kognition als Informationsverarbeitung viel weiter gefasst und haben nicht unbedingt etwas mit bewussten Erfahrungen zu tun. Dennoch ist es sehr plausibel, dass bei der als Informationsverarbeitung verstandenen Kognition Fehler auftreten, die dann zu kognitiven Fehlleistungen führen, die zumindest gelegentlich auf der Verarbeitung von Trägern falscher Informationen beruhen. Dass gelegentlich Träger von Fehlinformationen auftreten ist insbesondere deshalb wahrscheinlich, weil diese Träger das Produkt biologischer Prozesse sind, die niemals hundertprozentig ihre Funktion erfüllen²⁴.

In Übereinstimmung mit den hier vorgetragenen Ergänzungen zur Rechtfertigung der Annahme, dass Missrepräsentation möglich sein muss, schreibt Bill Ramsey:

„For example, when theorists posit inner representations, they typically assume that they now have an important way to explain how the system can fail to behave appropriately. It is now possible to explain faulty behavior as sometimes stemming from false representations of the world. In fact, considerable philosophical effort has been devoted to explaining how it is actually possible for a physical state to be in error – to misrepresent the nature of reality. This is an important topic because the possibility of misrepresentation is built into our ordinary way of understanding what it is to represent. If someone announced that they were using a technical notion of representation that didn't admit of misrepresentation, we would not think that this is just another way of handling the problem of error. Instead, we would think that whatever the posited state was doing, it wasn't playing a representational role. We can't posit representational states to do many of the things they are supposed to do in a theory unless the posit itself is sufficiently similar to the sort of things we pre-theoretically think representations are.“ (Ramsey 2007, 32)

Es ist also durchaus angebracht auch in Bezug auf subpersonale Prozesse der Informationsverarbeitung an der Annahme, dass Missrepräsentation möglich sein muss, festzuhalten (Vgl. Godfrey-Smith 2007, 107; Deacon 2010, 148).

²³ Husserl hätte von „Erlebnissen“ gesprochen. Erlebnisse sind der zentrale Untersuchungsgegenstand von Husserl Phänomenologie. Unter Erlebnissen versteht Husserl „diejenigen Vorkommnisse, zu denen ich einen epistemisch privilegierten Zugang habe: Ein Ereignis ist genau dann eines meiner Erlebnisse, wenn es Eigenschaften hat, von denen gilt: ich kann introspektiv mit Evidenz urteilen (ich kann „adäquat wahrnehmen“), daß sie ihm zukommen.“ (Künne 1986, 174).

²⁴ Hierbei setze ich voraus, dass zumindest einige unsere Sinnesorgane die Funktion haben bestimmte Umweltbedingungen korrekt zu repräsentieren.

1.3 Das Problem der Missrepräsentation

1.3.1 Das Problem der Missrepräsentation in naturalistischen Ansätzen

In diesem Abschnitt soll kurz verdeutlicht werden, inwiefern die Bedingung, dass Missrepräsentationen möglich sein müssen, für naturalistische Theorien des Gehalts eine echte Herausforderung darstellt. Das Problem besteht allgemein gesprochen darin, dass naheliegende Antworten darauf, wie man den Gehalt naturalistisch festlegen könnte, Missrepräsentationen unmöglich machen.

„[...] naturalistic theories of mental content have the unintended consequence of making error and misrepresentation impossible.“ (Perlman 2000, xvii)

Machen wir uns dies an zwei Beispielen klar, an einer einfachen kausalen Theorie²⁵ und an Dretskes informationstheoretischem Ansatz (Dretske 1981).

Eine einfache Idee den Gehalt einer Repräsentation naturalistisch zu bestimmen, wäre es zu sagen, dass der Gehalt einer Repräsentation immer die Ursache der Repräsentation ist.

"The unifying inspiration for causal theories of mental content is that some syntactic item "X" means X because "X"s are caused by Xs." (Adams & Aizawa 2010)

Diese einfache kausale Theorie ist zumindest im Bereich der Wahrnehmung prima facie plausibel; eine Repräsentation, die durch das Auftauchen eines Fressfeindes verursacht wird, hat den Gehalt *Fressfeind* (*im Gesichtsfeld aufgetaucht*); eine Repräsentation, die durch das Auftauchen von Beute verursacht wird, hat den Gehalt *Beute* (*im Gesichtsfeld aufgetaucht*) usw. Das Problem ist nun aber, dass in dieser einfachen kausalen Theorie Missrepräsentationen unmöglich sind. Es kann nicht sein, dass eine Repräsentation den Gehalt *Fressfeind* hat, ohne dass diese tatsächlich von einem Fressfeind verursacht wurde. Was auch immer die Ursache der Repräsentation ist, ist zugleich auch der Gehalt der Repräsentation. Ist die Ursache der Repräsentation eine Fressfeind-Attrappe, dann wäre der Gehalt *Fressfeind-Attrappe* und nicht etwa *Fressfeind*: Missrepräsentation ist somit unmöglich.

Kommen wir nun zum zweiten Beispiel. Fred Dretske hat 1981 in *Knowledge and the Flow of Information* einen Ansatz entwickelt, der nicht auf der Kausalrelation basiert, sondern auf einer informationalen Relation, die er aus der Informationstheorie von Shannon entwickelte.

Dretskes Grundidee in (1981) ist grob gesagt Folgende: Eine (mentale) Repräsentation *r* hat den Gehalt *p*, wenn *r* die Information, dass *p* trägt und ein Signal *r* trägt die Information, dass *p*, wenn die bedingte Wahrscheinlichkeit, dass *p* der Fall ist, wenn *r* auftritt bei 1 (bei 100%) liegt (Vgl. Beckermann 2008, 359f). Unter Voraussetzung dieser Idee kann es aber wieder keine Missrepräsentation geben. Wenn die Repräsentation *r* die Information, dass *p* trägt, dann ist *p* eine notwendige Bedingung für *r*. Folglich ist es unmöglich, dass eine Repräsentation *r* den Gehalt *p* hat, wenn *p* nicht der Fall ist (Vgl. Beckermann 2008, 360f).

Wir können also festhalten, dass die Bedingung, dass Missrepräsentationen möglich sein müssen, für naturalistische Theorien des Gehalts tatsächlich eine Herausforderung darstellt (Vgl. auch Neander 2012 und Godfrey-Smith 1989). Die soeben beschriebene Schwierigkeit wird in der Literatur oft auch kurz als Problem der Missrepräsentation bezeichnet (Vgl. Rakova 2006, 118).

²⁵ Einen guten Überblick über aktuelle kausale Theorien des Gehalts bieten Adams & Aizawa (2010).

1.3.2 Das Disjunktionsproblem

In diesem Abschnitt soll kurz das sogenannte Disjunktionsproblem vorgestellt werden, weil das Problem der Missrepräsentation in der Literatur vor allem im Zusammenhang mit dem Disjunktionsproblem debattiert wurde und wird (für eine Übersicht siehe Neander 2012; Detel 2001, 609-616; Prinz 2002, 243-249²⁶). Aufgrund der engen Verbindung, die zwischen dem Problem der Missrepräsentation und dem Disjunktionsproblem gesehen wird (Dretske 1986, 28-32; Godfrey-Smith 1989, 534; Fodor 1990, 59-61; Neander 2012; Wild 2010a, 370), entsteht zudem der Eindruck, dass eine Lösung des Problems der Missrepräsentation nicht ohne Lösung des Disjunktionsproblems möglich ist²⁷.

Das Disjunktionsproblem wurde 1984 von Jerry Fodor in „Semantics, Wisconsin Style“ eingeführt. Bei meiner Darstellung orientiere ich mich aber an Fodor (1990).

Wir können zunächst wieder von unserer einfachen kausalen Theorie ausgehen, die besagte, dass der Gehalt einer Repräsentation immer die Ursache der Repräsentation ist. Fodors Formulierung lautet:

„Suppose, to put it crudely, that „dog“ means *dog* (and thus has dogs and only dogs in its extension) because it's a law that dogs cause „dogs“.“ (Fodor 1990, 59)

Durch die gesetzmäßige Verbindung, die Fodor anspricht, wird zudem deutlich, dass es darum geht, den Gehalt von Repräsentationstypen zu bestimmen.

Nach Fodor gibt es nun zwei Möglichkeiten:

- 1) „Hund“ wird nur durch Hunde verursacht. Daraus folgt aber, dass „Hund“ nicht falsch werden kann²⁸. Mit anderen Worten: Missrepräsentation ist unmöglich.
- 2) „Hund“ wird auch durch Nicht-Hunde (z.B. Katzen bei Nacht²⁹) verursacht. Aufgrund unserer Bestimmung des Gehalts durch die Kausalrelation folgt daraus, dass der Gehalt von „Hund“ nun eine Disjunktion ist, nämlich *Hunde oder Katzen-bei-Nacht*³⁰. D.h. die Extension von „Hund“ sind Hunde und Katzen bei Nacht. Daraus wiederum folgt, dass eine „Hund“-Repräsentation, die von einer Katze bei Nacht verursacht wurde, wahr ist. Mit anderen Worten: Missrepräsentation ist unmöglich.

Das Disjunktionsproblem besteht also darin, dass eine Repräsentation nicht durch etwas verursacht werden kann, was nicht zur Extension der Repräsentation gehört. Dies wiederum ist unmöglich, weil die Extension alles umfasst, was die Repräsentation verursacht. Wir werden später an verschiedenen Stellen auf das Disjunktionsproblem zurückkommen.

²⁶ Prinz konzentriert sich allerdings auf Fodors Lösung des Disjunktionsproblems.

²⁷ Dieser Eindruck täuscht meines Erachtens. Wir werden später sehen warum (siehe Abschnitt 2.2.3.2).

²⁸ Vielleicht wäre es hier angemessener von „inkorrekt“ oder „unzutreffend“ zu reden als von „falsch“.

²⁹ „being (the right sort of) cat-on-a-dark-night“ (Fodor 1990, 59).

³⁰ Daher auch der Name Disjunktionsproblem.

1.3.3 Missrepräsentationsverständnis

In diesem Abschnitt soll angegeben werden, was in dieser Arbeit als ein paradigmatischer Fall von Missrepräsentation gelten soll. Auf diesen Fall werden wir uns dann in der Arbeit konzentrieren. Die Einschränkung der Betrachtungen auf einen bestimmten Fall schwächt natürlich die Verallgemeinerbarkeit der Resultate, die wir eventuell finden, ab. Sie hat aber den grossen Vorteil, dass zumindest einigermaßen klar ist, wovon wir sprechen, wenn von Missrepräsentation die Rede ist. Hinzu kommt, dass es im Rahmen dieser Arbeit aufgrund der Vielzahl von Phänomenen und Situationen, bei denen vermutlich Missrepräsentationen eine Rolle spielen, auch aus praktischen Gründen nötig ist eine Fokussierung vorzunehmen.

In einem ersten Schritt wollen wir einige Phänomene von den Betrachtungen ausschliessen, die auch intuitiv nicht unmittelbar als Missrepräsentationen involvierend verstanden werden. Hierzu zähle ich Phänomene, bei denen Repräsentationen von Fiktivem, von Zukünftigem, von Kontrafaktischem, von Abstraktem eine Rolle spielen.³¹ Diese Phänomene stehen insofern im Zusammenhang mit dem Problem der Missrepräsentation als die informational bzw. kausalen Theorien, die wir bereits kurz betrachtet haben, nicht nur Missrepräsentationen, sondern auch die oben genannten Repräsentationen unmöglich machen (Vgl. Scholz 2009, 90-93).

Ebenso ausschliessen wollen wir Phänomene, die man als Missrepräsentationen bezeichnen könnte, wie Lügen bzw. absichtliches Täuschen³² oder die Verwendung von leeren Begriffen (Phlogiston).

Stattdessen wollen wir uns auf Wahrnehmungssituationen, die (zwingend) eine Reaktion erfordern, konzentrieren. Die Flucht eines Organismus' vor einem Fressfeind, nachdem er den Fressfeind wahrgenommen hat, ist eine solche Situation. Im Rahmen dieser Arbeit gehen wir davon aus, dass die Fluchtreaktion des Organismus' eine kognitive Leistung ist. Weiterhin gehen wir davon aus, dass die Leistung zum Teil auf der Verarbeitung von Informationen beruht und dass in diesem Verarbeitungsprozess Informationsträger eine Rolle spielen, die die Anwesenheit des Fressfeindes repräsentieren bzw. deren Gehalt man in etwa mit *Fressfeind anwesend* wiedergeben kann. Nehmen wir an, wir hätten eine naturalistische Theorie des Gehalts entwickelt, d.h. wir haben u.a. einen Weg gefunden, ganz allgemein Typen von Gehaltsträgern einen bestimmten Gehalt zuzuordnen. Nehmen wir weiterhin an, dass es uns mit Hilfe unserer Theorie und empirischen Untersuchungen gelungen ist, den Typ von Informationsträgern zu identifizieren, der die Anwesenheit von Fressfeinden repräsentiert. Nun ist folgende Situation denkbar:

Es ist kein Fressfeind anwesend, dennoch wird ein Informationsträger produziert³³, der die Anwesenheit von Fressfeinden repräsentiert, und der Organismus ergreift die Flucht.

Der in dieser Situation produzierte Informationsträger zeigt die Anwesenheit von einem Fressfeind an, obwohl keiner da ist. Dieser Informationsträger ist somit falsch, er ist eine Missrepräsentation.

Wir wollen in dieser Arbeit die soeben dargestellte Situation als den paradigmatischen Fall von Missrepräsentation betrachten und uns bei unseren Überlegungen auf diesen Fall beschränken.

³¹ Die Aufzählung beansprucht keine Vollständigkeit.

³² Interessanterweise behauptet Skym (2010, 73): "Any theory that says that deception is impossible is a non-starter". Damit hat er aber vermutlich nur insofern Recht, als Täuschung (deception) die Möglichkeit von Missrepräsentation voraussetzt.

³³ Wenn wir uns unter dem Informationsträger ein Muster von neuronalen Impulsen vorstellen, dann ist die Rede von "produzieren" nicht merkwürdig.

Analog zu der genannten Situation lassen sich Situationen denken, in denen Informationsträger die Anwesenheit von Beute anzeigen, ohne dass diese da ist. Solche und ähnliche Situationen, d.h. Situation, in denen etwas angezeigt wird, was tatsächlich nicht da ist, zählen auch zum paradigmatischen Fall von Missrepräsentation. Etwas anders gelagert sind Situationen, in denen die Beute oder Fressfeinde da sind, aber die entsprechenden Informationsträger, die deren Anwesenheit anzeigen, nicht produziert werden. Teilweise werden aber auch solche Situationen in unseren Überlegungen eine Rolle spielen. Die Beispiele werden sich in dieser Arbeit also weitestgehend auf Beutefang- und Fluchtverhalten beschränken.

Insgesamt sollten wir nun ein hinreichend genaues Verständnis davon haben, was wir in dieser Arbeit unter einer Missrepräsentation verstehen wollen.

1.3.4 Der teleosematische Ansatz

1.3.4.1 Die grundlegende Idee im Ausgang von Dretske

In diesem Abschnitt geht es darum, die grundlegende Idee des teleosemantischen Ansatzes zur Lösung des Problems der Missrepräsentation vorzustellen. Der teleosemantische Ansatz beruht darauf, die Funktion von Repräsentationen mit ins Spiel zu bringen, wobei sich die Möglichkeit zur Missrepräsentation nun aus der Möglichkeit zur Fehlfunktion ergeben soll (Vgl. Laurence and Margolis 2003, 270; Menary 2007, 103; McDonald & Papineau 2006, 4).

„The trouble, in effect, is that the simple causal account of representation leaves no room for misrepresentation. Teleosemantic theories promise to deal better with misrepresentation. Theories of this kind ask about the purpose of the perceptual concept, in a biological sense, rather than about its causes (with biological purposes cashed out aetiologically, in terms of histories of natural selection). The referential value of the concept can then be equated with those items which it is the biological function of the concept to track. Since concepts can malfunction, like other biological traits, it no longer follows that misrepresentation is impossible. Sometimes a concept will be activated when it is not supposed to be. (See Millikan 1984, 1989, Papineau 1984, 1993a.)“ (Papineau 2002, 113)

Alle Theorien, die durch diesen Zug charakterisiert sind, kann man als teleosemantische Theorien bezeichnen (Vgl. Millikan 2008, 97-107). Die grundlegende teleosemantische Idee wurde unabhängig voneinander von Millikan (1984) und Papineau (1984) entwickelt.

Seitdem Dretske die Probleme seines informationstheoretischen Ansatzes erkannt hat, vertritt er ebenfalls eine teleosemantische Theorie. Wir können uns an seiner Version die grundlegende Idee des teleosemantischen Lösungsansatzes klar machen. Diese lässt sich bei Dretske wie folgt wiedergeben: Ein Tachometer *R* hat die Funktion die Geschwindigkeit eines Wagens, z.B. 60 km/h, zu repräsentieren. Der Tachometer ist so entworfen, dass die Tachoanzeige *r* (Zeiger auf „60“) die Information *p* (Wagen fährt 60 km/h) trägt. Ist der Tachometer defekt, dann behält er trotzdem seine Funktion. Auf diese Weise ist es möglich, dass die Tachoanzeige *r* (Zeiger auf „60“) den Gehalt *p* (Wagen fährt 60 km/h) hat, auch wenn der Wagen nicht 60 km/h fährt (Dretske 2009, 385).

Um diese Idee auf mentale Repräsentationen übertragen zu können und den naturalistischen Ansprüchen gerecht zu werden, vertritt Dretske wie auch andere Teleosemantiker einen etiologischen Funktionsbegriff, nach dem die Funktion eines Dinges derjenige Effekt ist, für den es in einem Evolutionsprozess selektiert wurde. Der etiologische Funktionsbegriff ist also durch die Darwins Theorie der Natürlichen Selektion fundiert und gilt damit als naturwissenschaftlich respektabel, ausserdem beruht er nicht wie im Tachometer-Beispiel auf den bereits gegebenen Absichten von menschlichen Designern. Das Wesentliche am etiologischen Funktionsbegriff ist, dass er es erlauben soll zwischen dem Haben einer Funktion und dem tatsächlichen Ausüben einer Funktion unterscheiden zu können. In diesem Sinne sind etiologische Funktionen normativ.

"This appeal to functional role does two things for Dretske. First, it enables him to handle the problem of misrepresentation in an intuitively appealing way. Error is now possible because we can say that whenever an indicator is triggered by something other than what it is supposed to indicate – something other than what it was, in fact, selected to indicate – it is a case of misrepresentation." (Ramsey 2007, 130)

Damit dürfte die Grundidee des teleosemantischen Ansatzes hinreichend klar sein. Insgesamt lässt sich feststellen, dass der teleosemantische Lösungsansatz allgemein als wichtiger Fortschritt gegenüber rein kausalen, rein bildtheoretischen oder rein informationstheoretischen Ansätzen angesehen wird³⁴ und von Vielen auch als gelungene Lösung des Problems der Missrepräsentation (z.B. Neander 2012; Shea 2005, 21-30; MacDonald und Papineau 2006, 11-12; Wild 2010a, 19; Godfrey-Smith 2013, 51; Millikan 2013a, 105; Neander in Vorbereitung, Kapitel 4, 20-23).

³⁴ Eine Ausnahme zu dieser Auffassung bildet Jerry Fodor, z.B. in Fodor (1990).

1.3.4.2 Herausforderung: Die Unbestimmtheit der Funktion

In diesem Abschnitt soll kurz dargelegt werden, warum der teleosemantische Lösungsansatz – obwohl er vermutlich einen Fortschritt darstellt – nicht ohne Weiteres überzeugen kann. Dies gilt vor allem dann, wenn dem teleosemantischen Lösungsansatz kein ausgearbeiteter Funktionsbegriff zugrunde liegt. Ohne einen ausgearbeiteten Funktionsbegriff gerät nämlich die teleosemantische Lösung in Schwierigkeiten, die dem Disjunktionsproblem analog sind (Vgl. auch Abschnitt 3.2.1). Wir können uns dies an einem Beispiel von Fred Dretske klar machen, das mittlerweile ziemlich berühmt ist.

„Some marine bacteria have internal magnets (called magnetosomes) that function like compass needles, aligning themselves (and, as a result, the bacteria) parallel to earth magnetic field. Since these magnetic lines incline downwards (towards geomagnetic north) in the northern hemisphere (upwards in the southern hemisphere), bacteria in the northern hemisphere, oriented by their internal magnetosomes, propel themselves towards geomagnetic north. The survival value of magnetotaxis (as this sensory mechanism is called) is not obvious, but it is reasonable to suppose that it functions so as to enable the bacteria to avoid surface water. Since the organisms are capable of living only in the absence of oxygen, movements towards geomagnetic north will take the bacteria away from oxygen-rich surface water and towards the comparatively oxygen-free sediment at the bottom.“ (Dretske 1986, 26)

Nach Dretske scheint es vernünftig anzunehmen, dass die Magnetosomen die Funktion haben durch ihre Ausrichtung anzuzeigen, wo sich sauerstoffreiches Wasser befindet. Unter dieser Annahme wird eine Ausrichtung der Magnetosomen, die durch einen Stabmagneten beeinflusst, das Bakterium in tödliches sauerstoffreiches Oberflächenwasser leiten, zu einer Missrepräsentation (Vgl. Dretske 1986, 27). Mit anderen Worten: Ob eine Missrepräsentation vorliegt, hängt unmittelbar davon ab, wie wir zuvor die Funktion der Magnetosomen bestimmt haben. Diesen Zusammenhang und das damit verbundene Problem hat Dretske klar erkannt. Er schreibt:

„There is, however, a more serious objection to this approach to the problem of misrepresentation. Consider, once again, the bacteria. It was said that it was the function of their magnetotactic system to indicate the whereabouts of oxygen-free environments. But why describe the function of this system in this way? Why not say that it is the function of this system to indicate the direction of [...] magnetic (not necessarily geomagnetic) north. This primitive sensory mechanism is, after all, functioning perfectly well when, under the bar magnet's influence, it leads its possessor into a toxic environment. *Something* is going wrong in this case, of course, but I see no reason to blame the sensory mechanism, no reason to say it is not performing *its* function.“ (Dretske 1986, 29)

Das Problem ist analog zum Disjunktionsproblem, weil wir die Funktion der Magnetosomen prinzipiell auch disjunktiv bestimmen könnten, auch wenn dies hier biologisch weniger plausibel erscheint. Wir könnten sagen, dass es die Funktion der Magnetosomen ist, die Richtung von sauerstoffreichem Wasser anzuzeigen *oder* die Richtung, in der sich ein Stabmagnet befindet.

Insgesamt können wir also festhalten, dass der teleosemantische Lösungsansatz ohne einen gut ausgearbeiteten und biologisch fundierten Funktionsbegriff nicht überzeugend ist. Etwas allgemeiner formuliert: Der teleosemantische Lösungsansatz ist zwar vielversprechend, aber

die Idee allein, ohne eine ausgereifte Theorie, ist nicht ausreichend, um das Problem der Missrepräsentation zu lösen.

1.3.4.3 Die Einschätzung der teleosemantischen Lösung

In diesem Abschnitt soll kurz anhand einiger Zitate belegt werden, dass die Einschätzung des teleosemantischen Ansatzes zur Lösung des Problems der Missrepräsentation in der Literatur offen ist. D.h. die Einschätzung des teleosemantischen Ansatzes reicht von: a) gelungen über b) offen, umstritten bis hin zu c) verfehlt.

a) *Gelungen*

„The outlines of a physicalist account of representational content have been developed most fully and convincingly, to my mind, by Fred Dretske (1981, 1986, 1988). For other promising accounts, see Stampe 1977, Millikan 1984, Harman 1987, Fodor 1990, Proust 1997.“ (Kriegel 2009, 66)

„Equally, it seems it [a teleological theory] can, in principle, overcome the problem of misrepresentation.“ (Lowe 2000, 98)

b) *Offen, Umstritten*

"[...], it is crucial for a theory of representation to be able to account for misrepresentation and the possibility of error as well, since one of the central features of representations is that they have truth values and truth conditions. They can be true or false, and there are conditions under which they are true and others under which they are not. But this presents a problem for the causal theory. If x represents y if and only if x has been caused in the appropriate way by y, misrepresentation is pretty much ruled out. More sophisticated forms of causal theories have subsequently been developed in order to tackle these various difficulties, but so far there is no general agreement that they have succeeded in doing so." (Gallagher and Zahavi 2008, 112f)³⁵

"While representationalists hold that a naturalistic account of representation can be given, the project of actually giving such an account is far from finished, and there are significant disagreements about how to proceed. There are, however, several ideas in the field that look promising. Dretske (1988, 1995) and Millikan (1984) are uppermost in my thoughts in this paragraph, but these and similar ideas are widely known and accepted by many." (Robinson 2004, 54)

„Ein prominenter Ausweg ist die Erweiterung der kausalen zu einer ‚teleofunktionalistischen‘ Bedeutungstheorie, welche Repräsentationsgehalte auf Normalbedingungen relativiert und auf diese Weise zu spezifizieren sucht. Die Normal- oder Optimalbedingungen werden ihrerseits evolutionstheoretisch hergeleitet. Ob diese normativen Prädikate sich tatsächlich auf naturalistisch akzeptable Weise erläutern lassen, ist umstritten. Soweit ich sehe, ist es den naturalistischen Bedeutungstheorien bisher nicht gelungen, die behauptete enge Verwandtschaft von

³⁵ Gallagher und Zahavi fassen Millikans Theorie als verfeinerte Form eines kausalen Ansatzes auf.

konventioneller Bedeutung und natürlichen Anzeichen plausibel zu machen.“ (Keil 2003, 93)³⁶

c) Verfehlt

„ET [Etiological teleosemantics] accounts of content are not vague and non-reductive, but they deliver the wrong contents.“ (Mendola 2008, 109)

„We will see why they [teleological/adaptional theories, AW] actually are use theories and why they fail to make room for misrepresentation.“ (Perlman 2000, xix)

Bisher konnten also weder die Befürworter noch die Gegner des teleosemantischen Ansatzes die Sache eindeutig für sich entscheiden. Damit bleibt insbesondere auch die Frage offen, ob die Auffassung von Kognition als Informationsverarbeitung durch eine teleosemantisch ausgerichtete Theorie des Gehalts den Homunkulus-Einwand vermeiden kann.

³⁶ Keils Einschätzung ist besonders interessant, weil Keil (1993b) Millikans Biosemantik noch als degenerierendes Forschungsprogramm bezeichnet hat (siehe auch Abschnitt 3.2.3).

1.3.5 Anspruch an die Lösung des Problems

In diesem Abschnitt wird erläutert, an welchem Anspruch in dieser Arbeit Millikans Lösungsansatz zum Problem der Missrepräsentation gemessen und geprüft werden soll. Kurz vorweggenommen: Es geht nur darum herauszufinden, ob Millikans Ansatz eine mögliche Lösung ist, und noch nicht darum, ob die Lösung des Ansatzes auch die tatsächlich in der Natur realisierte Lösung ist.

Wir hatten im Abschnitt zum Theorieverständnis (S. 15) gesagt, dass eine Theorie aus zwei Gründen scheitern kann:

1. die Theorie ist inkonsistent
2. die Theorie hält einer Überprüfung an der Realität nicht stand

Zudem wollen wir die Bedingung, dass Missrepräsentation möglich sein muss, als ein Postulat in jede naturalistische Theorie des Gehalts aufnehmen.

Wir können demzufolge eine Lösung des Problems der Missrepräsentation so verstehen, dass zwischen der Umsetzung der beiden Ziele, nämlich eine allgemeine Bestimmung davon, was ein Gehaltsträger ist (Ziel 1) und wie deren Gehalt festgelegt wird (Ziel 2) anzugeben und der Bedingung, dass Missrepräsentation möglich sein muss, kein Widerspruch besteht. Es geht also bei der Lösung des Problems der Missrepräsentation zumindest im Rahmen dieser Arbeit primär um Konsistenz und nicht schon um die Frage, ob die betreffende Theorie wahr ist.

Ich unterscheide in dieser Arbeit also zwischen:

- 1) wahren/gut bestätigten Theorien
- 2) möglichen/konsistenten Theorien
- 3) unmöglichen/inkonsistenten Theorien

Dies ist vermutlich erläuterungsbedürftig. Zur dritten Kategorie sollen all diejenigen Theorien gehören, die nicht konsistent sind und deshalb scheitern. Alle übrigen Theorien gehören in die zweite Kategorie und einige davon auch zur ersten. Unter den möglichen theoretischen Vorstellungen ist dann noch empirisch zu prüfen, welche davon tatsächlich in der Natur realisiert sind. Die realisierten gehören zu 1), die nicht realisierten zu 2).

Die kausalen oder informationstheoretischen Gehaltstheorien, die wir hier betrachtet haben, würden also demnach zur dritten Kategorie gehören. Dies ist so, weil sich aus der theoriespezifischen Umsetzung des Zieles 2, d.h. der allgemeinen Angabe, darüber, wodurch der Gehalt von Gehaltsträgern festgelegt ist, ableiten lässt, dass Missrepräsentationen unmöglich sind, was im Widerspruch zum Postulat der Möglichkeit von Missrepräsentationen steht. So gesehen sind diese Theorien nicht konsistent.

Den Unterschied zwischen möglichen und wahren Theorien können wir uns an einem Beispiel aus der Molekularbiologie verdeutlichen.

Die meisten Lebensvorgänge beruhen letztlich auf der spezifischen Funktionalität von Enzymen. Diese ergibt sich aus ihrer räumlichen Struktur, die wiederum das Resultat einer ganz spezifischen Anordnung von 20 verschiedenen Aminosäuren ist. Wie kann man nun erklären, dass Organismen über Generationen hinweg immer wieder exakt die gleichen Aminosäuresequenzen synthetisieren und damit die lebensnotwendigen Enzyme? Nachdem der Aufbau der DNA entdeckt war (Watson und Crick 1953), versuchte sich George Gamow an einer solchen Erklärung (Vgl. Wunder 2008, 16-22). Seiner Theorie zufolge lagern sich die Aminosäuren direkt an Hohlräume der spiraliggewundenen DNA-Doppelhelix an. Die

Hohlräume werden durch vier Basen bestimmt, wobei zwei der vier Basen ein komplementäres Paar bilden (Siehe Abbildung 3).

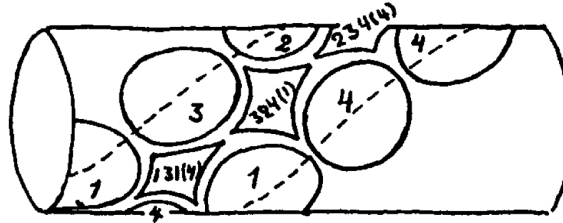


Abbildung 3: Schematische Darstellung der DNA, in der die spezifischen Zwischenräume eingezeichnet sind. Die Ziffern 1 bis 4 stehen für die vier Basen, wobei 1 und 2 sowie 3 und 4 ein komplementäres Paar bilden. (George Gamow, "Possible Relation between Deoxyribonucleic Acid and Protein Structures", *Nature* 173 (1954), 318.)

Dadurch ist jeder Hohlraum aufgrund der Komplementarität von Adenin und Thymin sowie von Cytosin und Guanin durch drei Basen festgelegt. Insgesamt ergibt sich auf diese Weise eine Übersicht, die zwanzig verschiedene Zwischenräume angibt (Siehe Abbildung 4).

1 1 2 a.	2 1 2 b.	3 1 2 c.	4 1 2 d.
3 3 4 e.	2 3 4 f.	3 3 4 g.	4 3 4 h.
1 1 2 i.	1 1 2 j.	1 1 2 k.	2 1 2 l.
1 1 2 m.	1 1 2 n.	3 3 4 o.	3 3 4 p.
3 3 4 q.	3 3 4 r.	3 3 4 s.	4 3 4 t.

Abbildung 4: Die zwanzig verschiedenen Zwischenräume auf einen Blick. Jeder Zwischenraum ist mit einem Kleinbuchstaben gekennzeichnet. (George Gamow, "Possible Relation between Deoxyribonucleic Acid and Protein Structures", *Nature* 173 (1954), 318.)

Gamow hat also eine mögliche Lösung dafür geliefert, wie die spezifische Sequenz der Enzyme in der DNA konserviert wird: Durch eine Art Schlüssel-Schloss-Prinzip könnte die DNA genau 20 Aminosäuren codieren. Gamows Theorie ist testbar³⁷, und seine Lösung hat sich nicht als die tatsächlich in der Natur realisierte erwiesen. Heute wissen wir, dass eine andere mögliche Lösung, nämlich die der sogenannten „Adaptor-Hypothese“ tatsächlich in der Natur realisiert ist. Nach der Vorstellung dieser Hypothese sollte es Moleküle geben, die sowohl an Aminosäuren als auch an Nukleinsäuren binden können, und daher als Adaptor fungieren (Vgl. Crick 1955, 8), wobei sich später herausstellt hat, dass das Adaptor-Molekül eine Sorte von RNA ist, die sogenannte Transfer-RNA (t-RNA).

³⁷ Bestimmte Rauten (z.B. Raute a und Raute h) können nicht nebeneinander stehen. Folglich dürften auch einige Aminosäurepaarungen in den Enzymen nicht auftreten. Derartige Einschränkungen wurden empirisch jedoch nicht bestätigt.

Uns sollte nun der Unterschied zwischen möglichen Lösungen und tatsächlich realisierten Lösungen verständlich sein. Zudem sollte klar sein, dass man die Frage, ob eine bestimmte Theorie bzw. Hypothese eine mögliche Lösung für ein bestimmtes Problem darstellt, prinzipiell unabhängig von der Frage beantworten kann, ob die Lösung auch tatsächlich in der Natur realisiert ist.

In diesem Sinne soll es innerhalb dieser Arbeit nur darum gehen, ob ein bestimmter teleosemantischer Lösungsansatz, nämlich derjenige von Ruth Millikan, eine mögliche Lösung des Problems der Missrepräsentation darstellt oder ob dies nicht der Fall ist.

Nun könnte man meinen, dass der Anspruch, nur eine mögliche Lösung für das Problem der Missrepräsentation anzubieten, gering ist. Dies relativiert sich jedoch, wenn wir das Problem als durch eine sogenannte Wie-ist-es-bloß-möglich-Frage (how-possible question) motiviert verstehen³⁸. Wie-ist-es-bloß-möglich-Fragen implizieren im Gegensatz zu einfachen Wie-Fragen (how questions), deren Beantwortung auf die Schilderung eines Ablaufes oder die Angabe einer von mehreren Alternativen hinausläuft, ein zumindest prima facie unüberwindliches Hindernis (obstacle) aufgrund dessen etwas, p, zunächst unmöglich erscheint, so dass man sich fragt: wie ist p bloß möglich? (Vgl. Cassam 2007; Dray 1966) Viele philosophische Probleme beruhen auf solchen Wie-ist-es-bloß-möglich-Fragen. So zum Beispiel: Wie ist es möglich, dass wir einen freien Willen haben (wenn doch alle Handlungen kausal determiniert sind)? (Vgl. Nozick 1981, 8) Mögliche Antworten bestehen dann im Beiseitigen des Hindernisses, etwa im Aufzeigen der Kompatibilität von freien Willen und Determinismus oder in einer Widerlegung des Determinismus. In diesem Sinne besteht eine mögliche Lösung des Problems der Missrepräsentation darin zu zeigen, dass Missrepräsentationen trotz der Annahme einer naturalistischen Theorie des Gehalts (obstacle) möglich sind.

Die zentrale Frage dieser Arbeit ist, ob Millikans Biosemantik eine solche mögliche Lösung bereitstellt. Die Frage ist nicht und kann nicht sein, ob Millikans naturalistische Theorie des Gehalts wahr ist. Hiefür wären grundsätzlich empirische Untersuchungen nötig, die im Rahmen dieser Arbeit nicht geleistet werden können. Hinzu kommt, dass man für eine solche empirische Untersuchung zunächst einen Weg finden müsste die repräsentationalistische Theorie des Geistes empirisch überprüfbar zu machen³⁹.

³⁸ Zum Nutzen solcher Fragen vgl. auch Nichols und Stich (2003, 60).

³⁹ Vgl. Abschnitt 1.2.2, insbesondere die Schwierigkeit bei der empirischen Überprüfung, die sich ergibt, wenn die Indikatorhypothesen ihrerseits nicht gesichert sind.

1.4 Ziel der Arbeit

1.4.1 Ziel, Vorgehen, Bedeutung

In diesem Abschnitt geht es darum, a) das Ziel der Arbeit zu formulieren, b) kurz zu erläutern, wie das Ziel erreicht werden soll und c) welche Bedeutung das Resultat der Arbeit hat.

a) Ziel

Das Ziel dieser Arbeit ist die Beantwortung der Frage, ob Millikans Theorie des Gehalts, die Biosemantik, das Problem der Missrepräsentation löst oder ob ihr dies nicht gelingt.

b) Vorgehen

Die Beantwortung dieser Frage erfordert eine eingehende Untersuchung von Millikans Biosemantik. Innerhalb der Untersuchung gehe ich davon aus, dass die Biosemantik auf zweierlei Weise am Problem der Missrepräsentation scheitern kann:

- 1) die Ziele einer Theorie des Gehalts – die allgemeine Bestimmung davon, was ein Gehaltsträger ist (Ziel 1) und die allgemeine Angabe, darüber, wodurch der Gehalt von Gehaltsträgern festgelegt ist (Ziel 2) – werden tatsächlich naturalistisch umgesetzt, aber die Möglichkeit der Missrepräsentation entpuppt sich unter diesen Bedingungen als eine lediglich scheinbare
- 2) Missrepräsentation ist innerhalb der Theorie tatsächlich möglich, aber die Umsetzung der Ziele einer Theorie des Gehalts entpuppt sich als lediglich scheinbar naturalistisch⁴⁰

Bevor also der kritische, prüfende Teil der Arbeit begonnen werden kann, ist zunächst eine fundierte Darstellung davon nötig,

- i) was nach Millikan ein Gehaltsträger ist (Ziel 1)
- ii) wie nach Millikan der Gehalt eines Gehaltsträgers allgemein bestimmt werden kann (Ziel 2)
- iii) wie nach Millikan innerhalb ihrer Theorie Missrepräsentation möglich sein soll.

An den darstellenden Teil schließt sich die kritische Prüfung der Theorie an, die aus der Rekonstruktion und Bewertung von Argumenten gegen Millikans Lösung und eigenen Einwänden besteht. Hierbei beanspruche ich keine Vollständigkeit, aber doch die zentralen Argumente der Diskussion aufgenommen zu haben. Auf der Basis dieser Prüfung soll dann geurteilt werden, ob Millikan das Problem der Missrepräsentation löst oder ob ihr dies nicht gelingt.

c) Bedeutung

Das Urteil dieser Untersuchung über Millikans Lösung des Problems der Missrepräsentation kann natürlich nur vorläufigen Charakter haben. Es ist prinzipiell immer möglich, neue, bisher übersehene, aber entscheidende Einwände vorzubringen oder bisher für gültig gehaltene Einwände durch neue Argumente zu entkräften.

Die Bedeutung des Resultats dieser Untersuchung soll hier nur Hinblick auf seine Bedeutung für die Auffassung von Kognition als Informationsverarbeitung erläutert werden.

⁴⁰ Natürlich scheitert die Biosemantik auch, wenn sie beide Fehler – Missrepräsentation unmöglich (1) und nicht naturalistisch (2) – aufweist.

Wenn sich herausstellen sollte, dass Millikans Ansatz keine konsistente Theorie (Vgl. Abschnitt 1.2.2) ist und es zudem keine alternativen naturalistischen Gehaltstheorien gibt, die mögliche Lösungen darstellen, dann wäre die Informationsverarbeitungsauffassung von Kognition aufgrund des Homunkulus-Einwandes diskreditiert.

Sollte sich hingegen herausstellen, dass Millikans Theorie eine mögliche Lösung ist, dann wäre auch Informationsverarbeitungsauffassung von Kognition konsistent zumindest in Hinblick auf den Homunkulus-Einwand.

Aus einem positiven Resultat folgt aber nicht, dass die Informationsverarbeitungsauffassung von Kognition die beste theoretische Grundannahme zur Erklärung von kognitiven Fähigkeiten ist. Insbesondere folgt nicht, dass die Annahme von Repräsentationen, Gehalts- oder Informationsträgern überhaupt zur Erklärung von kognitiven Fähigkeiten notwendig sind (Vgl. dazu z.B. die Argumentation von Stich 1983).

Durch die eingehende und bisher noch nicht durchgeführte Untersuchung von Millikans Lösung zum Missrepräsentationsproblem liefert die Arbeit einen kleinen aber nicht unwichtigen Puzzlestein zur Bewertung des Informationsparadigmas in den Kognitionswissenschaften. Das Informationsparadigma ist zwar nicht unumstritten, aber an seiner Aktualität und Relevanz kann dennoch kein Zweifel bestehen. So geht beispielsweise eines der grössten EU-Forschungsprojekte, das Human-Brain-Projekt (HBP), explizit von der Informationsverarbeitungsauffassung der Kognition aus.

„The HBP should thus include a concerted programme of theoretical research, focusing on the mathematical principles underlying the relationships between different levels of brain organisation and the way the brain acquires, represents and stores information.“
(HBP Report 2012, 11)⁴¹

Eine adäquate Einschätzung der Informationsverarbeitungsauffassung der Kognition ist aufgrund der Relevanz dieses Paradigmas für so viele Forschungsbereiche (KI-Forschung, Robotik, Kognitive Psychologie, Neurowissenschaften etc.) besonders dringend und wichtig. Die vorliegende Arbeit versucht hierzu einen kleinen Beitrag zu leisten.

⁴¹ “The human brain is a massively complex information processing system” (HBP Report 2012, 17).

1.4.1 Warum Millikans Lösung?

In diesem Abschnitt soll kurz begründet werden, warum in dieser Arbeit Millikans Lösungsansatz zum Missrepräsentationsproblem untersucht wird und nicht die Lösung einer anderen naturalistischen Theorie des Gehalts, etwa die von Dretske (1981, 1988), Fodor (1990), Papineau (1984), Neander (forthcoming), Skyrms (2010) oder Bischof (1998).

Bevor man aber die Wahl dieser oder jener Theorie begründet, könnte man sich fragen, warum nicht gleich alle derzeit vorhandenen Theorien dieser Art untersucht wurden. Dies wäre natürlich wünschenswert gewesen, weil dann die Resultate der Untersuchung eine gewisse Allgemeingültigkeit beanspruchen könnten. Eine Untersuchung solchen Umfangs war aber im Rahmen dieses Forschungsprojektes aus Zeitgründen nicht möglich. In diesem Zusammenhang spielt auch meine Überzeugung eine Rolle, dass eine Überprüfung und Bewertung der jeweiligen Theorie nur auf der Basis von gründlichen Kenntnissen der jeweiligen Theorie vorgenommen werden kann. Dies gilt insbesondere für Millikans Theorie, die eine erhebliche Komplexität aufweist (Vgl. Macdonald 1995: 248; Kurthen 1992: 147, 181; Lenk 2001: 235; Beckermann 2008: 368), so dass bereits die gründliche Darstellung der Theorie einen beträchtlichen Teil dieser Arbeit einnimmt. Eine Untersuchung mehrerer Theorien wäre im Rahmen dieses Forschungsprojektes also nur auf Kosten der notwendigen Gründlichkeit möglich gewesen, weswegen ich mich dagegen entschieden habe.

Kommen wir nun zu der Frage: Warum Millikans Theorie? Dazu ist prinzipiell zu sagen, dass es für die hier entwickelte Problematik gleichermassen sinnvoll gewesen wäre eine der anderen Theorien zu untersuchen, dennoch sprechen einige Gründe für die Wahl von Millikans Theorie.

Der Hauptgrund ist, dass Millikans Theorie die am weitesten ausgearbeitete naturalistische Theorie des Gehalts ist.⁴²

„Ruth Millikan's account is by far the most carefully articulated, bristling with implications about the details of the other philosophical approaches to meaning mentioned above.“ (Dennett 1995, 403)

Aufgrund des hohen Ausarbeitungsgrades von Millikans Theorie lassen sich in dieser Theorie auch Antworten auf Detailfragen finden, die in weniger ausgearbeiteten Theorien unbeantwortet bleiben müssen. Millikan skizziert nicht nur einen Ansatz, sondern arbeitete ihre Theorie über Jahre hinweg aus. Millikans Theorie gilt daher Vielen als aussichtsreichster Versuch einer naturalistischen Bedeutungstheorie (Vgl. Detel 2001b: 468) und es ergibt Sinn sich als erstes den aussichtsreichsten Kandidaten anzuschauen.

Der hohe Ausarbeitungsgrad von Millikans Theorie hat auch zu einer gewissen Komplexität der Theorie geführt, die eine ausführliche Darstellung nötig macht. Eine solche Darstellung liegt für Millikans Theorie in der hier erzielten Ausführlichkeit nicht vor und kann seit langem als Desiderat in den Debatten um naturalistische Theorie des Gehalts gelten⁴³. Durch die Wahl von Millikans Theorie kann diese Forschungslücke geschlossen werden.

Ein weiterer Grund für die Untersuchung von Millikans Theorie ist, dass Millikan (explizit) auch das erste Ziel einer naturalistischen Theorie des Gehalts verfolgt, nämlich zu bestimmen, was überhaupt ein Gehaltsträger, eine Repräsentation, ist (Vgl. Millikans „A General Theory of Signs“ in Millikan 1984, Part II).

⁴² Neanders Theorie (Neander forthcoming) könnte Millikan in diesem Punkt Konkurrenz machen. Ähnliches gilt m.E. für Norbert Bischofs Theorie (Bischof 1998, 273-363), die aber wenig Beachtung gefunden hat, was vermutlich daran liegt, dass es bislang keine englische Übersetzung des Textes gibt.

⁴³ Umfangreichere Darstellungen von Millikans Theorie finden sich in Detel (2001b), Wild (2010a), Shea (2005).

Millikan entwickelt zudem einen ausgefeilten Funktionsbegriff (Vgl. Millikans „Theory of Proper Functions“ in Millikan 1984, Part II). Ein solcher Funktionsbegriff ist, wie wir bereits erläutert haben (Vgl. Abschnitt 1.3.4.2), für die teleosemantische Lösung des Disjunktionsproblems von besonderer Bedeutung. Weder Dretske noch Papineau, die ebenfalls teleosemantische Ansätze vertreten, haben einen ähnlich gut ausgearbeiteten Funktionsbegriff⁴⁴.

Darüber hinaus lässt sich der später folgenden Darstellung vorgreifend feststellen, dass die Biosemantik genaugenommen zwei Lösungen für das Problem der Missrepräsentation anbietet. Die erste basiert auf missgebildeten Funktionsträgern, die zweite auf abNormalen (terminus technicus von Millikan) Umweltbedingungen. Dies wird nicht ausreichend beachtet, was dazu führt, dass die Ressourcen der Biosemantik zur Lösung des Problems der Missrepräsentation unterschätzt werden. Es ist daher ein wichtiger Beitrag meiner Arbeit, diese beiden Lösungen klar auseinanderzuhalten und getrennt zu bewerten.

Schließlich gehört Millikans Ansatz neben denen von Fodor und Dretske zu den einflussreichsten Arbeiten auf dem Gebiet naturalistischer Theorien des Gehalts. Es ist daher sinnvoll die Untersuchung solcher Theorien bei ihren einflussreichsten Vertretern zu beginnen.

⁴⁴ Neanders Funktionsbegriff ist ebenfalls gut ausgearbeitet (Vgl. Neander 1991).

1.4.2 Millikans Anspruch

In diesem Abschnitt geht es darum klarzustellen, dass der Anspruch, den Millikan mit der Biosemantik verbindet, deutlich über das hinausgeht, was wir von der Biosemantik erwarten. Fragen, die sich aus Millikans höherem Anspruch ergeben, werden in dieser Arbeit nicht behandelt.

Wir erwarten von der Biosemantik lediglich zweierlei:

- 1) dass sie eine allgemeine Bestimmung davon liefert, was ein Gehaltsträger ist, und allgemein angibt, wodurch der Gehalt von Gehaltsträgern festgelegt ist, ohne dabei geistbegabte Wesen bzw. deren Fähigkeiten vorauszusetzen und
- 2) dass sie dies auf eine Weise tut, die Missrepräsentationen nicht unmöglich macht.

Insgesamt – so ist zumindest die Hoffnung – soll dies im Rahmen des Informationsverarbeitungsparadigmas dazu beitragen auch geistige Fähigkeiten, wie Denken und Sprechen, erklären zu können.

Millikan hingegen beansprucht, dass die Biosemantik auch eine adäquate Theorie für Gehaltsträger wie Gedanken und Sätze ist, wobei für Millikan Gedanken und Sätze durchaus Eigenschaften haben, die nicht allen Gehaltsträgern zukommen (Vgl. Millikan 1989b, 294-297). Millikan beansprucht also für die Biosemantik einen wesentlich grösseren Anwendungsbereich als er für diese Arbeit nötig ist und hier angenommen werden soll. Mit ihrem Anspruch wird aber auch der Verstehensaspekt bzw. die Sinnenebene von Gehaltsträgern wieder relevant, die in dieser Untersuchung ausgeklammert werden sollen (Vgl. Abschnitt 1.2.4). Millikans Strategie mit der Sinnenebene umzugehen besteht darin, den Verstehensaspekt als am wenigsten wichtig, als nahezu irrelevant hinzustellen (Vgl. Millikan 1984, 11). Hinter dieser Strategie verbirgt sich Millikans Überzeugung, dass unsere bisherige Auffassung über Gedanken und Sätze mit einer grundlegend fehlgeleiteten Position verbunden sind (Vgl. Millikan 1984, 157). Diese vermeintlich irriige Position bezeichnet Millikan als Bedeutungsrationismus (meaning rationalism). Der Bedeutungsrationismus ist nach Millikan eine Konjunktion von drei erkenntnistheoretischen Thesen:

„1. *The epistemic givenness of meaning identity and difference* A rational person has the capacity to discern a priori whether or not any two of her thoughts comprehend the same term or proposition, the same meaning.

2. *The epistemic givenness of univocity* A rational person has the capacity to discern a priori when she is entertaining a thought with double or ambiguous meaning [...].

3. *The epistemic givenness of meaningfulness* A rational person has the capacity to discern a priori whether she is meaning a term or proposition or whether her thought is empty of meaning.“ (Millikan 1993e, 287)

Millikan zufolge durchzieht der Bedeutungsrationismus die gesamte Philosophische Tradition.

„But in one form or another, meaning rationalism permeates nearly every nook and cranny of our philosophical tradition.“ (Millikan 1984, 92)

„Meaning rationalism is not a single doctrine but a syndrom.“ (Millikan 1984, 326)

Das Minimalziel in *Language, Thought, and other Biological Categories* (1984), Millikans Hauptwerk, ist sogar in Bezug auf den Bedeutungsrationismus bestimmt und es besteht darin, den Bedeutungsrationismus durch den Kontrast einer Alternative zumindest sichtbar und damit einer kritischen Hinterfragung zugänglich zu machen.

„In order even to to comprehend what meaning rationalism is, what various forms it can take, it is necessary forcefully to fling down on the table something with which to contrast it. The description of intentionality that I will offer is designed to serve that purpose. I also think that it is true. But if nothing else comes of it, may it at least leave behind this question in a relatively clear form: what if any grounds do we have for continuing blindly to bet on meaning rationalism, given its monotonous no-show track record.“ (Millikan 1984, 92)

Vielleicht kann man sogar ohne Übertreibung sagen, dass das Hauptanliegen von Millikans Schaffen die Bekämpfung des Bedeutungsrationismus ist. Dass dieses Anliegen für Millikan eine wahre Herzensangelegenheit ist, belegen die folgenden beiden Zitate.

„My desire is to kill meaning rationalism dead, and then beat on it.“ (Millikan 1993a, 12)

„I am driven by a naive but fanatical desire to make it utterly clear not only to Alice but to Every Reader that, and why, meaning rationalism is a Very Bad Idea and Hopeless Dead End.“ (Millikan 1993e, 303)

Aus Millikans Anspruch an die Biosemantik ergeben sich etliche Fragen: „Ist der Bedeutungsrationismus tatsächlich falsch?“, „Wenn ja, durchzieht er tatsächlich fast die gesamte Philosophie?“, „Wenn ja, bietet Millikans Auffassung tatsächlich eine gangbare Alternative?“, etc. Diese und ähnliche Fragen werden im Rahmen der vorliegenden Arbeit keine Berücksichtigung finden, da hier von einem geringeren Anspruch an die Biosemantik ausgegangen wird.

1.5 Methode

Die vorliegende Arbeit richtet sich nicht nach einer spezifischen Methode, wie etwa der Begriffsanalyse oder der Hermeneutik. Dennoch möchte ich auf drei methodologische Punkte eingehen, an denen ich mich orientiert in dieser Arbeit habe.

Der *erste methodologische Punkt* betrifft die Darstellung von Millikans Lösung. Diese Darstellung ist durch Überlegungen geleitet, die eine gewisse Ähnlichkeit mit Davidsons „principle of charity“ aufweisen.

„Die grundlegende methodologische Vorschrift besagt daher, dass eine gute Interpretationstheorie Einigkeit maximiert.“ (Davidson 1990, 244)

Mir geht es bei der Darstellung von Millikans Lösung weniger um Textgenauigkeit und Exegese, sondern darum Millikans Lösung möglichst stark und konsistent zu präsentieren. Diese Maximierungs- bzw. Optimierungsabsicht bezüglich ihrer Lösung erlaubt es mir gewisse Ansichten und Aussagen von Millikan zu ignorieren. Es kann also vorkommen, dass meine Darstellung mit einigen Aussagen in Millikans Texten nicht übereinstimmt. Natürlich darf die Anzahl der nicht übereinstimmenden Aussagen, ganz analog zu Davidsons Prinzip, eine kritische Masse nicht überschreiten. Ich würde aber darüber hinaus auch eine Interpretation bevorzugen, die nicht die Einigkeit in Millikans Aussagen maximiert, sondern die Stärke und Konsistenz ihrer Lösung, zumindest solange, wie sich die Lösung überhaupt noch Millikan zuschreiben liesse⁴⁵. Mein Vorgehen weicht hierin also von Davidsons Prinzip ab. In diesem Punkt wird auch deutlich, dass diese Arbeit primär eine systematische und keine historisch- exegetische ist. In dieser systematischen Absicht liegt zugleich die Rechtfertigung für dieses Vorgehen.

Der *zweite methodologische Punkt* betrifft meinen intensiven Einsatz von Abbildungen und Übersichten in der gesamten Arbeit. Dieser Einsatz gründet schlicht darin, dass mir bildliche Darstellungen beim Verstehen komplexer Zusammenhänge helfen. Natürlich gehe ich davon aus, dass es zumindest einigen Lesern ebenso geht. Die Bilder liefern dabei eine Konkretisierung und Zusammenfassung auf Kosten der Möglichkeit, Dinge abstrakt und allgemein zu halten, wie dies in einem reinen Text möglich ist. In dieser Vereinfachung liegt vermutlich auch einer der grossen Vorteile einer bildlichen Darstellung (Vgl. dazu Larkin, J. H. und H. A. Simon 1987).

Der *dritte methodologische Punkt* betrifft die Rekonstruktion der Einwände gegen Millikans Lösung des Problems der Missrepräsentation. Hierbei habe ich mich wesentlich an den Kapiteln 6 und 8 von Holm Tetens' Einführung *Philosophisches Argumentieren* (2006) orientiert. Insbesondere habe ich versucht die Einwände gegen Millikans Lösung in einer Form darzustellen, die Tetens als Standardform bezeichnet. Die Standardform ist eine Darstellung bzw. Rekonstruktion des Argumentes, in der sich alle Prämissen und die Konklusion auf Anhieb erkennen lassen (Vgl. Tetens 2006, 59). Diese Form der Darstellung ist für die Beurteilung eines Argumentes sehr hilfreich. Tetens geht sogar noch einen Schritt weiter:

„Ein Argument hat nur verstanden, wer es prinzipiell in der Standardform reformulieren kann. Diese These lässt sich einfach begründen. Wer argumentiert, führt die problematische Wahrheit einer Aussage auf die unproblematische oder schon

⁴⁵ Die Einschränkung, dass es sich noch erkennbar um Millikans Lösung handeln muss, verhindert, dass ich eine Konsistenzmaximierung vornehme, die den dritten Teil der Arbeit, die kritische Prüfung von Millikans Lösung, überflüssig macht. Es geht mir darum, dass Millikans Theorie in ihrer stärksten Interpretation geprüft wird, auch wenn diese Interpretation nicht diejenige ist, die den stärksten Rückhalt in Millikans Texten finden würde.

gesicherte Wahrheit anderer Aussagen zurück. Diese Rückführung gelingt, falls das Argument schlüssig ist. Schlüssigkeit ist zwar nicht die einzige Eigenschaft, die ein gutes Argument aufweisen muss, aber sie ist wiederum nicht irgendeine Eigenschaft von Argumenten unter «ferner liefern». Wer die Schlüssigkeit eines Arguments nicht zu beurteilen vermag, dem entgeht etwas Wesentliches an ihm. So wie Argumente normalerweise vorgetragen werden, gerät man leicht in Zweifel, ob sie schlüssig sind. Wer sich dessen vergewissern will, muss Argumente nach Prämissen und Konklusion ordnen können. Ausserdem ist die Folgerungsbeziehung zwischen Prämissen und Konklusion möglichst transparent zu formulieren. Für beides, für die Ordnung der Sätze nach Prämissen und Konklusion und für eine durchsichtige Formulierung des inferentiellen Zusammenhangs zwischen den Sätzen, empfiehlt sich die Standardform.“ (Tetens 2006, 61-62).

Teil 2:

Millikans Lösung

2.1 Millikans Theorie des Gehalts

2.1.1 Millikans Zeichenbegriff

2.1.1.1 Millikans Zeichenbegriff: Die Grundidee

In diesem Abschnitt soll es darum gehen eine erste Vorstellung von Millikans Zeichenbegriff zu bekommen. Dies soll anhand eines stark vereinfachten, fiktiven Beispiels geschehen:

Für einen fiktiven im Wasser lebenden Organismus sind Fische Fressfeinde und Krebstiere Nahrung. Von Fischen schwimmt der Organismus weg und auf Krebstiere zu. Zu dieser (kognitiven) Leistung ist der Organismus in der Lage – so die Annahme bei Millikan – weil er über ein repräsentationales System verfügt. Das repräsentationale System ist dabei im Rahmen eines Sender-Empfänger-Modells zu verstehen (siehe Abbildung 5).

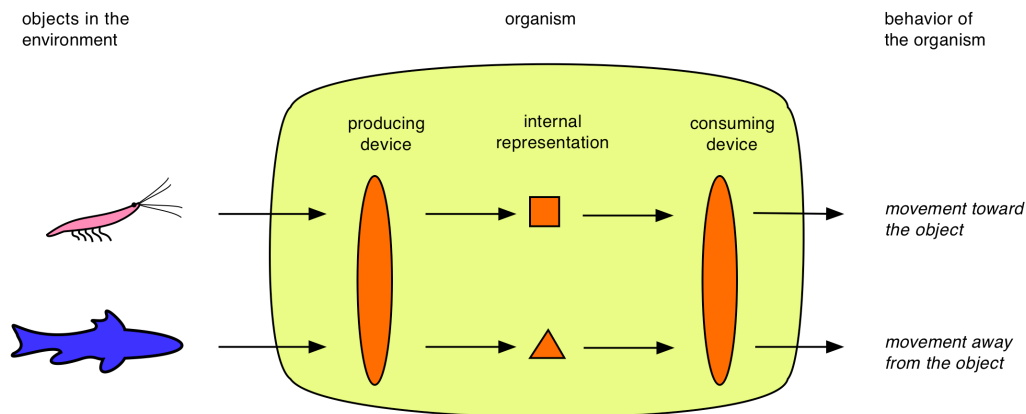


Abbildung 5: Repräsentationales System eines fiktiven Organismus

Der Sender des repräsentationalen Systems⁴⁶ ist das sensorische System des Organismus. Millikan bezeichnet den Sender als producing device. Der producing device hat die Funktion in Abhängigkeit von bestimmten Umweltbedingungen, Fische oder Krebstiere anwesend, etwas, einen bestimmten Botenstoff oder einen bestimmten Nervenimpuls, zu produzieren. Dies geschieht nach folgender Regel:

- Wenn Krebstiere anwesend, dann Nervenimpuls A (durch Viereck symbolisiert)
- Wenn Fische anwesend, dann Nervenimpuls B (durch Dreieck symbolisiert)

Der Empfänger des repräsentationalen Systems ist das effektorische System des Organismus. Millikan bezeichnet den Empfänger als consuming device. Der consuming device hat die Funktion in Abhängigkeit von dem, was der producing device produziert hat, nämlich einen bestimmten Botenstoff oder einen bestimmten Nervenimpuls, ein bestimmtes Verhalten, hin- oder wegschwimmen, auszulösen. Dies geschieht wiederum nach einer Regel:

- Wenn Nervenimpuls A (durch Viereck symbolisiert), dann hinschwimmen
- Wenn Nervenimpuls B (durch Dreieck symbolisiert), dann wegschwimmen

⁴⁶ Das repräsentationale System ist der Sender und der Empfänger zusammen.

Der producing device und der consuming device sind also so aneinander angepasst, dass sich der Organismus durch ihr Zusammenspiel in der jeweiligen Umweltsituation vorteilhaft verhält.

Das, was der Sender in einem solchen Zusammenhang produziert und der Empfänger konsumiert, ist nach Millikan ein Zeichen.

2.1.1.2 Millikans Definition eines Zeichens

In diesem Abschnitt soll damit begonnen werden, unser Verständnis von Millikans Zeichenbegriff zu vertiefen. Hierzu werde ich zunächst Millikans Zeichendefinition angeben und alle technischen Ausdrücke der Definition identifizieren. Das Verständnis dieser technischen Ausdrücke ist unerlässlich, um Millikans Zeichenbegriff vollständig zu erfassen und um ihre Lösung des Problems der Missrepräsentation bewerten zu können.

Ruth Millikan gibt folgende Definition von „Zeichen“ bzw. „intentional icon“⁴⁷:

„(1) A sentence is a member of a reproductively established family having direct proper functions. [...]

(2) Normally⁴⁸ a sentence stands midway between two cooperating devices, a producer device and an interpreter device, which are designed or standardized to fit one another, the presence and cooperation of each being a Normal condition for the proper performance of the other.

(3) Normally the sentence serves to adapt the cooperating interpreter device to conditions such that proper functions of that device can be performed under those conditions.

(4a) In the case of imperative sentences, it is a proper function of the interpreter device, as adapted by the sentence, to produce conditions onto which the sentence will map in accordance with a specific mapping function of a kind to be described below.

(4b) In the case of indicative sentences, the Normal explanation of how the sentence adapts the interpreter device such that it can perform its proper functions makes reference to the fact that the sentence maps conditions in the world in accordance with a specific mapping function of a kind to be described below.“ (Millikan 1984, 96f)

Bevor wir die technischen Ausdrücke identifizieren, sind drei Bemerkungen angebracht.

Bemerkung 1: Auch wenn in der Definition stets von Sätzen die Rede ist, wird hier dennoch „intentional icon“ definiert und nicht „Satz“.

„These four characteristics define the paradigm of what I will call an "intentional icon".“ (Millikan 1984, 97)

Für Millikan geht es aber zusätzlich darum zu zeigen, dass auch Sätze die vier genannten Bedingungen erfüllen⁴⁹. Dies hängt mit ihrem Anspruch zusammen (Vgl. Abschnitt 1.4.3) und ist für unsere Untersuchung irrelevant.

Bemerkung 2: Es gibt intentional icons, die sowohl indikativ als auch imperativ sind (Vgl. Millikan 1984, 99). Dies würde auch für die beiden Zeichen in unserem vereinfachten, fiktiven Beispiel gelten (Vgl. Abschnitt 2.1.1.1). Das „Viereckszeichen“ im Beispiel besagt demnach etwas, wie „Fressfeind anwesend (indikativ), wegschwimmen (imperativ)“. Später wird Millikan diese intentional icons als Pushmi-Pullyu-Repräsentationen bezeichnen (Vgl. Millikan 1995). Pushmi-Pullyu-Repräsentationen sind für die vorliegende Arbeit von

⁴⁷ Genaugenommen sind Zeichen und intentional icons bei Millikan nicht dasselbe: "There are a great many things in the world that are intentional icons. And all of the things that we are tempted to call "signs" exhibit very striking analogies to intentional icons. Hence I believe that the category of intentional icons is an important one to demarcate." (Millikan 1984, 96)

⁴⁸ "Normally" is eine Abkürzung für "when performing its proper functions in accordance with a Normal explanation" (Vgl. Millikan 1984, 96).

⁴⁹ "There are four characteristics that an indicative or imperative sentence Normally has that contribute to its being an intentional icon." (Millikan 1984, 96)

besonderem Interesse, denn sie sind nach Millikan „die bei weitem häufigsten *innerhalb* von Organismen vorkommenden intentionalen Zeichen“ (Millikan 2008 [2004], 220). Daher werden wir in dieser Untersuchung vor allem diesem Typ von intentionalen Zeichen berücksichtigen.

Bemerkung 3: Die Definition ist nicht so zu verstehen, dass alles, was diese Definition nicht erfüllt, kein Zeichen ist. Die Definition liefert vielmehr das Paradigma eines Zeichens. Alle anderen Zeichen sind „Zeichen“, weil sie ein oder mehrere wichtige Merkmale mit dem Paradigma teilen.

„There are a great many things in the world that are intentional icons. And all of the things that we are tempted to call "signs" exhibit very striking analogies to intentional icons.“ (Millikan 1984, 96)

Kommen wir nun zu den technischen Ausdrücken in der Definition, die erklärt werden müssen. Es sind dies:

- reproductively established family (reproduktiv etablierte Familie)
- direct proper function (direkte Eigenfunktion)
- Normal explanation (Normale Erklärung)
- Normal condition (Normale Bedingung)
- mapping function (Abbildungsfunktion)

Ebenso muss erklärt werden, was standardizing proper functions (standardisierende Eigenfunktionen) und relational proper functions (relationale Eigenfunktionen) sind, weil sonst die Bedingungen 2 und 3 der Zeichendefinition nicht korrekt verstanden werden können.

In den nächsten Abschnitten werden wir die folgenden Begriffe in der folgenden Reihenfolge erklären:

- 1) reproduction (Reproduktion)⁵⁰
- 2) first order reproductively established family (reproduktiv etablierte Familie erster Ordnung)
- 3) direct proper function (direkte Eigenfunktion)
- 4) Normal explanation (Normale Erklärung)
- 5) Normal condition (Normale Bedingung)
- 6) higher order reproductively established family (reproduktiv etablierte Familie höherer Ordnung)
- 7) relational proper function (relationale Eigenfunktionen)
- 8) standardizing proper functions (standardisierende Eigenfunktionen)
- 9) mapping function (Abbildungsfunktion)

Insbesondere die Begriffe 1-6 bauen dabei in folgender Weise auf einander auf:

⁵⁰ Millikans Definition von “reproductively established family” baut auf der Definition von “reproduction” auf (siehe Abbilung 6), so dass diese zuvor behandelt werden muss.

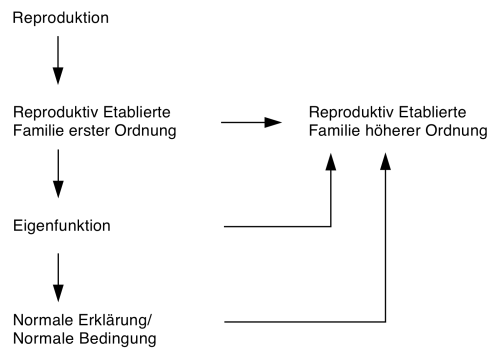


Abbildung 6: Abhängigkeit technischer Begriffe bei Millikan

Abschliessend sei noch angemerkt, dass es Millikan bei der Definition bzw. Konstruktion von Begriffen, wie „proper function“ oder „reproduction“ nicht darum geht eine Begriffsanalyse zu betreiben:

„My program is far removed from conceptual analysis; I need a term that will do a certain job, and so I must fashion one.“ (Millikan 1984, 19)

Eine der wenigen mir bekannten Stellen, in denen Millikan explizit auf ihre eigene Methode eingeht, ist die Folgende:

„Instead, I will attempt to construct a theory using another traditional philosophical method—by sifting and organizing certain matters of common knowledge, mixed with other observations easily made and with some information from contemporary science, then forming an explanatory hypothesis in an attempt to explain the data. In this case, the hypothesis will be about what work certain linguistic forms, certain kinds of words and phrases, manage to do and how and why they are able to do it. As philosophers usually do, I will construct this theory from the armchair, but in the way that the theoretical sciences also do, by reflecting on certain empirical data in an effort to understand them. How language works, I take it, is a broadly empirical question, to be investigated like any other.“ (Millikan 2010, 51f)

2.1.1.3 Erläuterung technischer Ausdrücke

2.1.1.3.1 Reproduction

In diesem Abschnitt soll Millikans Definition von Reproduktion eingeführt und erläutert werden. Millikans Definition von Reproduktion in leicht abgewandelter Form lautet:

B ist eine „Reproduktion“ von *A* gdw:

1. *B* einige bestimmte Eigenschaften p_1, p_2, p_3 etc. mit *A* gemeinsam hat und 2. erfüllt ist.
2. durch Natur- bzw. Spezialgesetze, die 3. Erfüllen, erklärt werden kann, warum *B* und *A* die Eigenschaften p_1, p_2, p_3 etc. gemeinsam haben
3. die Gesetze dies so erklären, dass die Kausalität von *A* nach *B* geht, d.h. hätte *A* bzgl. der Determinante x die Determinante, d.i. die Ausprägung p_1^* statt p_1 gehabt, dann hätte auch *B* p_1^* statt p_1 .⁵¹

Die Eigenschaften p_1, p_2, p_3 nennt Millikan reproduktiv etablierte Eigenschaften. Die Gesamtheit dieser Eigenschaften bezeichnet Millikan auch als reproduktiv etablierten Charakter und das „Original“ von dem die Reproduktion abstammt als Modell (model). D.h., wenn *B* die Reproduktion von *A* ist, dann ist *A* das Modell von *B*. Das Verhältnis von Determinante und Determinable erläutert Millikan wie folgt: Eine Eigenschaft p_1 ist bestimmt, relativ zu einer bestimmbaren Eigenschaft P , unter die p_1 und andere p_i ausschließende Eigenschaften fallen. *Rot* ist bestimmt, relativ zu *farbig*. Unter *farbig* fällt sowohl *rot* als auch *rot* ausschließende Eigenschaften wie *grün, gelb* etc. (Vgl. Millikan 1984, 19f).

Determinable		Determinante
farbig	→	rot, gelb, grün
rot	→	magenta, bordeaux
Stabilität von X	→	konkrete Messwerte x, y, z

Unter einem Spezialgesetz versteht Millikan ein Gesetz, das sich aus Naturgesetzen durch die Hinzufügung von konkreten Randbedingungen ableiten lässt⁵².

Als Beispiele für Reproduktionen nennt Millikan Gene, Viren⁵³, Begrüßungsgesten, Güter der Massenproduktion (wie Autos), Fußabdrücke, handwerkliche Produkte (ausser

⁵¹ Engl. Original: „An individual *B* is a ‚reproduktion‘ of an individual *A* iff:

- (1) *B* has some determinate properties p_1, p_2, p_3 , etc., in common with *A* and (2) below is satisfied.
- (2) That *A* and *B* have the properties p_1, p_2, p_3 , etc., in common can be explained by a natural law or laws operative in situ, which laws satisfy (3) below.
- (3) For each property p_1, p_2, p_3 , etc., the laws in situ that explain why *B* is like *A* in respect to p are laws that correlate a specifiable range of determinates under a determinable under which p falls, such that whatever determinate characterizes *A* must also characterize *B*, the direction of causality being straight from *A* to *B*.“ (Millikan 1984, 19f.).

⁵² So soll sich aus den Gesetzen der Physik und der Chemie sowie einer vollständig in physikalischen Begriffen verfassten Beschreibung einer Kopiermaschine ein Gesetz ableiten lassen, das die Konfiguration von Markierungen auf der Kopiervorlage (Original) mit den Markierungen auf der dazugehörigen Kopie korreliert (Vgl. Millikan 1984, 20).

Neuerfindungen) und sprachliche Ausdrücke. Verhalten, Organe oder Lebewesen sind hingegen keine Reproduktionen, denn deren Ähnlichkeit erklärt sich über die Ähnlichkeit von Genen.

„Rather, the stickleback’s *genes* and my *genes* (tokens) were reproduction of earlier gene tokens harbored by our respective parents, and similar *genes* produce similar *products*.“ (Millikan 1984, 21)

In diesem Umstand liegt auch der Grund weswegen Millikan zwischen reproduktiv etablierten Familien erster und höherer Ordnung unterscheidet.

Für ein genaues Verständnis von Millikans Reproduktionsbegriff scheint es mir besonders wichtig die zwei folgenden Fragen zu beantworten:

1. Was ist mit „Erklären durch ein Gesetz“ gemeint?
2. Ist eine Reproduktion immer so etwas wie eine Kopie im Sinne eines Duplikats?

Was ist also damit gemeint, dass das Haben der gleichen Eigenschaften von A und B (siehe Definition) durch ein Gesetz erklärt werden kann? Wie erklären Gesetze etwas?

Die Betonung der Rolle von Gesetzen für die Erklärung legt die Vermutung nahe, dass Millikan hier das deduktiv-nomologische Modell einer wissenschaftlichen Erklärung im Sinn hat. Bei einer deduktiv-nomologischen Erklärung (DN-Erklärung) wird der Satz, der das Explanandum-Phänomen (E) beschreibt, aus Sätzen, die die Antezedenzbedingungen ($A_1...A_k$) beschreiben, und Gesetzesaussagen ($G_1...G_r$), die insgesamt als Explanans bezeichnet werden, logisch abgeleitet (Vgl. Hempel 1977, 6)⁵⁴. Insgesamt hat die DN-Erklärung also die Gestalt eines Argumentes, wobei das Explanans den Prämissen und das Explanandum der Konklusion entspricht:

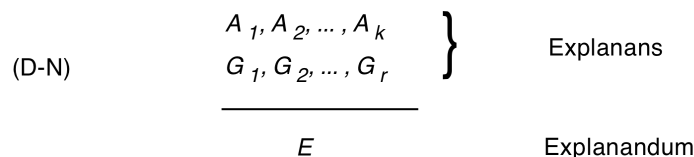


Abbildung 7: Das deduktiv-nomologische Erklärungsschema (in Anlehnung an Hempel 1977, 6)

Eine DN-Erklärung dafür, dass A und B dieselben Eigenschaften p_1, p_2, p_3 haben, könnte dann in verallgemeinerter Form in etwa so aussehen:

Explanans:

G: Befindet sich ein Objekt A vom Typ X mit der Merkmalsausprägung p_i unter der Bedingung C, dann durchläuft A eine Prozess P und es entsteht ein Objekt B vom Typ Y⁵⁵ mit derselben Merkmalsausprägung p_i .

A: A ist ein Objekt vom Typ X mit der Merkmalsausprägung p_i und befindet sich unter der Bedingung C.

Explanandum:

B ist vom Typ Y und hat die Merkmalsausprägung p_i (genau wie A).

⁵³ Viren sind m.E. keine Reproduktionen in Millikans Sinn, sondern müssen genauso wie Lebewesen betrachtet werden.

⁵⁴ Weil bei dieser Erklärung das zu erklärende Phänomen unter Gesetze subsumiert wird, spricht man auch vom *covering law model*.

⁵⁵ Typ X und Typ Y können identisch sein, müssen es aber nicht.

Die DN-Erklärung kann zwar schon lange nicht mehr als die generell akzeptierte Form einer wissenschaftlichen Erklärung gelten (Vgl. Salmon 1989), aber dies ist für Millikan unproblematisch. Für sie ist lediglich wichtig, dass es eine wissenschaftlich akzeptable Erklärung dafür gibt, dass *A* und *B* dieselben Eigenschaften p_1, p_2, p_3 haben und dabei die Kausalrichtung von *A* nach *B* verläuft. Welchen Typs eine solche Erklärung ist, ist dabei eher zweitrangig. Es könnte sich bei dieser Erklärung z.B. auch um eine mechanistische Erklärung im Sinne von Salmon (1989), Bechtel und Richardson (1993), Machamer et. al. (2000), Craver (2009) handeln.

Kommen wir nun zur zweiten Frage: Ist eine Reproduktion immer so etwas wie eine Kopie im Sinne eines Duplikats⁵⁶? Dieser Eindruck entsteht zumindest *prima facie* und wird von Millikan unterstützt, wenn sie meint, dass ihr Begriff der Reproduktion sehr nah an dem ist, was man in der alltäglichen Sprache als Kopie bezeichnen würde. Ebenso legen etliche von Millikans Beispielen (Viren, Gene, Güter der Massenproduktion) diesen Eindruck nahe. Dennoch ist er falsch. Eine Reproduktion kann eine Art Duplikat sein, muss es aber nicht. Andernfalls könnte nämlich ein Fußabdruck im Schlamm keine Reproduktion eines Fußes sein⁵⁷. Eine Reproduktion ist immer nur eine Reproduktion *relativ* zu einer oder mehreren Eigenschaften. Mit anderen Worten: Die Verwendung des Ausdrucks „Reproduktion“ im Sinne von Millikan erfordert zwingend auch eine Angabe der Eigenschaften hinsichtlich welcher der fragliche Gegenstand eine Reproduktion ist. Ohne die genaue Angabe der Eigenschaften kann sich hinter dem Ausdruck „ist eine Reproduktion des Gegenstandes A“ eine sehr heterogene Menge von Objekten verbergen, wie folgendes Bild veranschaulicht.

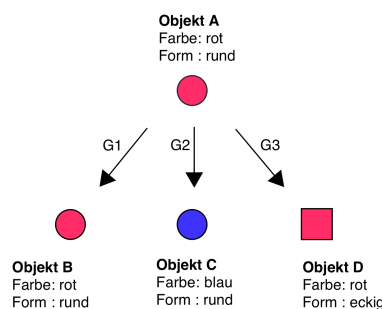


Abbildung 8: Objekt (A) und dessen Reproduktionen (B, C, D), die unterschiedliche Eigenschaften haben

Die Objekte B, C und D sind allesamt Reproduktionen von A, vorausgesetzt es gibt die entsprechenden Gesetze (G1, G2, G3) und dies gilt obwohl C und D keine der reproduzierten Eigenschaften gemeinsam haben. C und D sind zwar Reproduktionen von A, aber eben nicht hinsichtlich derselben Eigenschaften, C ist eine Reproduktion von A hinsichtlich der Form und D hinsichtlich Farbe.

Abschließend wollen wir ein konkretes Beispiel für eine Reproduktion im Sinne Millikans betrachten, auf das wir gelegentlich zurückkommen werden. Die RNA A hat die Basensequenz ACCGUUG⁵⁸ und aufgrund dessen eine bestimmte Resistenz gegen hydrolytischen Abbau

⁵⁶ Original und Reproduktion wären dann zwei kaum voneinander zu unterscheidende Objekte.

⁵⁷ Dies ist auch eines von Millikans Beispielen für eine Reproduktion (Vgl. Millikan 1984, 21).

⁵⁸ RNAs sind Makromoleküle, die aus einer Kette von Nukleotiden bestehen. Die vier Nukleotide der RNA sind: Adenosinphosphat (A), Guanosinphosphat (G), Cytidinphosphat (C), Uridinphosphat (U). Die spezifische polymere Verknüpfung der Nukleotide einer RNA wird auch als deren Basensequenz bezeichnet und kann durch die Angabe der Anfangsbuchstaben der verknüpften Nukleotide dargestellt werden.

(Eigenschaft p_1). Die RNA *B* mit derselben Basensequenz (ACCGUUG) und derselben Hydrolyseeigenschaft (Eigenschaft p_1) wie *A* ist eine Reproduktion der RNA *A*, wenn sie aufgrund doppelter basenkomplementärer⁵⁹ Anlagerung von freien Nukleotiden aus der RNA *A* hervorgegangen ist (siehe Abbildung 9), denn dieser Vorgang liefert eine gesetzmäßige Erklärung dafür, *warum* die RNA *B* die gleiche Hydrolyseeigenschaft (Eigenschaft p_1) wie die RNA *A* hat.

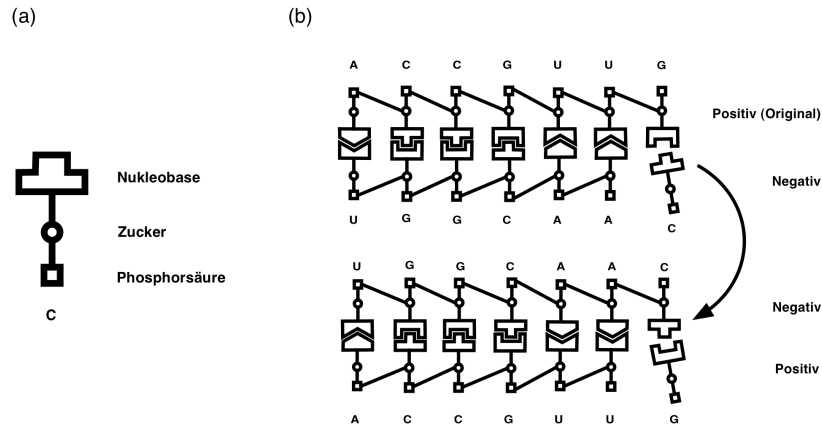


Abbildung 9: (a) Aufbau eines Nukleotids, (b) doppelte basenkomplementäre Anlagerung freier Nukleotiden an eine RNA (in Anlehnung an: Küppers 1983, 13 und 17)

⁵⁹ Zwischen jeweils zwei der Nukleobasen (siehe Abbildung 9) besteht eine sogenannte Basenkomplementarität. D.h.: „Zwischen Adenin und Uracil [...] einerseits und Guanin und Cytosin andererseits kommt es in wässriger Phase zur Ausbildung spezifischer Wasserstoffbrücken-Bindungen, wodurch sich sterisch nahezu identische Basenpaare bilden. Man sagt auch kurz: A und U [...] sowie G und C sind basenkomplementär zueinander.“ (Küppers 1986, 42)

2.1.1.3.2 First order reproductively established family

In diesem Abschnitt soll Millikans Definition von Reproduktiv Etablierter Familie erster Ordnung (FOREF) eingeführt und erläutert werden. Ihre Definition lautet:

Jede Gruppe von Gegenständen, die dieselben oder ähnliche reproduktiv etablierte Eigenschaften durch wiederholte Reproduktionen derselben Eigenschaft vom selben Modell oder Modellen erhalten hat, formen eine reproduktiv etablierte Familie erster Ordnung.⁶⁰

Ein genaueres Verständnis davon, was eine FOREF ausmacht, können wir erlangen, indem wir an einem fiktiven Beispiel versuchen die Mitglieder und damit den Umfang einer FOREF zu bestimmen.

In unserem Beispiel wollen wir von 18 RNA-Molekülen (siehe Abbildung 10) ausgehen, die alle dieselbe Resistenz gegen hydrolytischen Abbau haben (Vgl. S. 54). Einige der RNAs (1-4) haben sich spontan gebildet⁶¹, während die restlichen RNAs (5-18) durch doppelte basenkomplementäre Anlagerung (Vgl. S. 54f) entstanden sind.

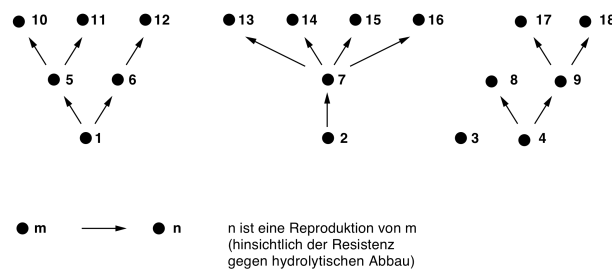


Abbildung 10: Fiktives Beispiel von RNAs, die teilweise Reproduktionen voneinander sind

Durch die Anwendung von Millikans Definition ergeben sich drei FOREFs (gelb unterlegte Flächen):

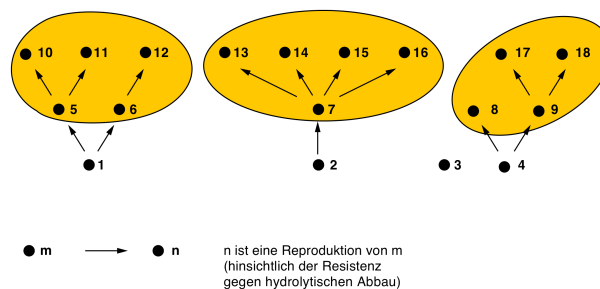


Abbildung 11: Fiktives Beispiel von RNAs, die drei FOREFs bilden

⁶⁰ Engl. Original: „Any set of entities having the same or similar reproductively established characters derived by repetitive reproductions from the same character of the same model or models form a *first-order reproductively established family*.“ (Millikan 1984, 23) Die Verwendung des Ausdrucks „reproduction“ in der Definition von FOREF entspricht nicht Millikans eigener Definition von „reproduction“. Nach ihrer eigenen Definition meint „reproduction“ ein Objekt, das bestimmten Kriterien entspricht, „reproduction“ in der Definition von FOREF hingegen meint einen Prozess dessen Resultat erst ein solches Objekt ist.

⁶¹ In einem geeigneten Medium und unter den richtigen Bedingungen können RNAs einfach spontan entstehen. „Liegen die monomeren Bausteine [die Nukleotide, A.W.] in energiereicher Form – zum Beispiel als Triphosphate – vor, so vereinigen sie sich auch ohne katalytische Unterstützung von Enzymen spontan zu makromolekularen Ketten [die Nukleinsäuren, A.W.].“ (Eigen 1992, 149).

Die RNAs 1-4 gehören keiner FOREF an, denn ihre spezifische Resistenz gegen hydrolytischen Abbau ist keine reproduktiv etablierte Eigenschaft. Für FOREFs gilt dasselbe wie für Reproduktionen, es gibt sie nur relativ zu einer bzw. mehreren Eigenschaften.

Dass die übrigen RNAs, deren Resistenz eine reproduktiv etablierte Eigenschaft ist, nicht eine grosse FOREF bilden, sondern drei verschiedene, liegt an der Bedingung, dass sie die Eigenschaft vom demselben Modell oder Modellen⁶² haben müssen. Millikan fasst hierbei den Begriff der Reproduktion und damit auch den des Modells transitiv auf.

„Trivially, if B is a reproduction of A and A is a reproduction of A' with respect to the same properties, then B is a reproduction of A'.“ (Millikan 1984, 20)

Insgesamt gilt also, dass alle Mitglieder einer FOREF mindestens ein Modell gemeinsam haben müssen, was darauf hinausläuft, dass sie alle von einem einzigen Modell abstammen müssen. Die Mitglieder einer FOREF sind somit gewissermaßen immer durch eine gemeinsame Kausalgeschichte verbunden.

Abschliessend soll kurz etwas dazu gesagt werden, dass nach Millikans Definition nicht nur die Weitergabe derselben Eigenschaften, sondern auch schon von ähnlichen ausreichen soll um eine FOREF zu formen. Dies scheint problematisch, weil die Definition an Präzision verliert. Dies liegt daran, dass Ähnlichkeit anders als Gleichheit eine graduelle Angelegenheit ist, was die Frage mit sich bringt, *wie* ähnlich sich die Reproduktion und ihr Modell hinsichtlich der relevanten Eigenschaft sein müssen. Ich werde diese Frage aber zunächst beiseite lassen und stattdessen versuchen zu ergründen, warum diese Aufweichung überhaupt nötig zu sein scheint.

Betrachten wir wieder unser Beispiel mit den RNAs. Die Komplementarität der Basen ist keine Garantie dafür, dass ausschließlich komplementäre Basen eine Bindung eingehen⁶³. Je länger eine RNA also ist, desto unwahrscheinlicher ist es, dass durch doppelte basenkomplementäre Anlagerung eine RNA mit exakt derselben Basensequenz und damit auch derselben Resistenz gegen hydrolytischen Abbau entsteht. Angenommen also aus einer RNA von grosser Kettenlänge entstehen durch doppelte basenkomplementäre Anlagerung weitere RNAs, dann hätte der grösste Teil dieser RNAs eine nahezu identische Basensequenz und folglich nur eine ähnliche Resistenz. Nur ein sehr geringer Prozentsatz hätte exakt dieselbe Sequenz und Resistenz und nur diese dürften wir bei einer strengeren Definition zu einer FOREF zusammenfassen. Das Motiv für den Einschub von „or similar“ in die Definition der FOREF könnte also darin liegen, interessante Gruppen von Individuen nicht grundlos aus unseren Betrachtungen auszuschließen. Uns würde dadurch etwas entgehen, nämlich, dass es hier eine Gruppe von Individuen gibt, die hinsichtlich einer Eigenschaft ziemlich homogen ist, wobei diese Homogenität auf kausal-historischen Zusammenhängen beruht. Solange der

⁶² Mit „oder Modellen“ ist nicht gemeint, dass die FOREF-Mitglieder entweder von einem oder von einem anderen Modell abstammen müssen (z.B. entweder von RNA 1 oder RNA 2). Der Einschub („oder Modellen“) ist vielmehr dem Umstand geschuldet, dass es nach Millikan auch Dinge gibt, die eine Reproduktion von verschiedenen Modellen zugleich sind. "It is possible for an item to have reproductively established properties derived from several different models. [...] an item can be a reproduction of numerous unlike things at once, deriving different properties from each. For example, a silver cross, such as many people wear on a chain around their necks, that is made also to resemble a vine of olive leaves has some reproductively established characters derived from earlier examples of Christian crosses and others derived from the olive leaves (or pictures of these) that the artist studied in order to make the cross." (Millikan 1984, 20f).

⁶³ „The process of molecular reproduction rests principally upon the affinity between two complementary nucleotides. Because the energies of interaction are finite and the molecules are always in thermal motion, incorrectly-matched pairs occur with a finite probability [...], so that in an ensemble of nucleic acids undergoing reproduction a certain fraction of the copies will always be defective.“ (Küppers 1983, 18)

Grad der Ähnlichkeit festgelegt ist, halte ich die Lockerung, die der Einschub von „oder ähnliche“ erzeugt, für gerechtfertigt und sinnvoll.

2.1.1.3.3 Proper Function

In diesem Abschnitt soll Millikans Definition von proper function (Eigenfunktion) eingeführt und erläutert werden. Millikans Grundidee ist dabei folgende:

Die Eigenschaften, die die Mitglieder einer reproduktiv etablierten Familie auszeichnen, können Effekte haben, die wiederum Einfluss auf die Entstehung weiterer Mitglieder nehmen. War dieser Einfluss in der Vergangenheit positiv, dann hat die betreffende Eigenschaft gleichsam für ihr eigenes fortwährendes Vorkommen gesorgt. Mitglieder einer reproduktiv etablierten Familie mit solchen Eigenschaften haben nach Millikan eine *Eigenfunktion* (*proper function*). Die Eigenfunktion ist der positive Effekt, den die Vorfahren dieser Mitglieder als Träger der reproduktiv etablierten Eigenschaften auf die Entstehung weiterer Mitglieder derselben reproduktiv etablierten Familie hatten.

Millikans genaue Definition von proper function (Eigenfunktion) lautet:

„Where *m* is a member of a reproductively established Family *R* and *R* has the reproductively established or Normal character *C*, *m* has the function *F* as a direct proper function iff:

1. Certain ancestors of *m* performed *F*
2. In part because there existed a direct causal connection between having character *C* and the performance of the function *F* in the case of those ancestors of *m*, *C* correlated positively with *F* over a certain set of items *S* which included these ancestors and other things not having *C*
3. One among the legitimate explanations that can be given of the fact that *m* exists makes reference to the fact that *C* correlated positively with *F* over *S*, either directly causing reproduction of *m* or explaining why *R* was proliferated and hence why *m* exists.“ (Millikan 1984, 96f)

Millikans Beispiele für Träger von Eigenfunktionen aus dem Bereich der Biologie sind Organe und Verhaltensmuster.

Zunächst fällt an der Definition von Eigenfunktion auf, dass der Begriff der Funktion bereits im Definiens auftaucht, d.h. die Definition von Eigenfunktion scheint bereits das Verständnis eines anderen Funktionsbegriffes vorauszusetzen. Auf diesen Punkt hat schon Henrik Walter hingewiesen (Vgl. Walter 1997, 192). Seiner Meinung nach ist dies aber unproblematisch, wenn der vorausgesetzte Funktionsbegriff seinerseits ergänzend definiert wird, z.B. in Anlehnung an Cummins (1975), und die beiden Funktionsbegriffe „sich nicht wechselseitig definieren“ (Vgl. Walter 1997, 192)⁶⁴. Ich bin mir nicht sicher, ob eine solche Ergänzung an dieser Stelle tatsächlich nötig ist. Anstatt von „hat die Funktion *F*“ könnte Millikan auch etwas umständlicher sagen „löst unter bestimmten Bedingungen den Effekt *E* aus“. In diesem Sinn induziert ein magnetisches Eisenstück, wenn man es durch eine Spule führt, in dieser Spule einen elektrischen Stromes. Mit anderen Worten: Millikan muss an dieser Stelle überhaupt kein Verständnis von Funktion voraussetzen, denn das Auslösen bzw. Verursachen von bestimmten Effekten unter bestimmten Bedingungen ist prima facie nichts, dass ausschließlich Dingen zukommt, die Funktionen im alltäglichen Sinn haben.⁶⁵

Schauen wir uns nun die zweite Bedingung genauer an. Sie fordert:

⁶⁴ Zudem muss der ergänzende Funktionsbegriff ebenfalls naturalistisch sein, was auf denjenigen von Cummins aber zu trifft.

⁶⁵ Insgesamt scheint mir aber Walters Hinweis gerechtfertigt, dass die Integration von verschiedenen Funktionen und damit die besondere Organisation des Lebendigen bei Millikan etwas zu kurz kommen und ergänzungsbedürftig scheinen.

a) die Eigenschaft C verursachte die Ausübung der Funktion F **UND** b) die Eigenschaft C korrelierte positiv mit der Ausübung der Funktion F⁶⁶

Die Glieder der Konjunktion sind logisch voneinander unabhängig, denn es gibt Korrelationen, die keine Verursachungen sind⁶⁷, und es gibt Verursachungen, die keine Korrelationen sind. Letzteres ist wahrscheinlich nicht ganz so offensichtlich. Ein Beispiel für eine Verursachung, die keine Korrelationen ist, liegt vor, wenn A als kausaler Faktor verstärkend auf B einwirkt, gleichzeitig jedoch A mit C korreliert, das als kausaler Faktor abschwächend auf B einwirkt, wobei As Verstärkung geringer ist als Cs Abschwächung. Wenn Rauchen Herzattacken verursacht, Rauchen aber gleichzeitig mit einer cholesterinarmen Diät korreliert ist, die die Gefahr von Herzattacken stärker abmindert als Rauchen diese hervorruft, dann verursacht Rauchen zwar Herzattacken, ist aber nicht mit diesen positiv korreliert. (Vgl. für Folgendes Sober 2000, 81) Kurzum: Millikan fordert beides, eine kausale Verbindung und eine positive Korrelation zwischen C und F.

Schauen wir uns nun die Bedingung 3 in der Definition von „Eigenfunktion“ genauer an. Ihr zufolge muss es unter den legitimen Erklärungen für die Existenz von m eine Erklärung geben, die sich explizit auf die positive Korrelation zwischen der reproduktiv etablierten Eigenschaft C und der Funktion F in einem Set S bezieht, um die Existenz von m zu erklären. Auch wenn es Millikan in der Bedingung 3 nicht explizit erwähnt, handelt es sich bei der gesuchten Art von Erklärung um eine Erklärung aus dem Gebiet der Evolutionsbiologie. Daran lassen ihre Bemerkungen und Ausführungen im Kontext der Definition von Eigenfunktion kaum Zweifel.

„But in the arena of evolutionary history, there is room for such statistical explanations. There are of course many legitimate explanations, some more interesting than others, for every happening in nature. What matters here is only that explanations making reference to correlations of a certain type can be given at all for why certain traits of organisms survive.“ (Millikan 1984, 26)

„Historical statistical correlations between organ or behavior structures and the performances or reliability of the performances of certain useful functions thus figure in legitimate explanations of why the modern Normal traits of any species have continued to proliferate thus culminating in the existence of current exemplars of these traits.“ (Millikan 1984, 27)

Die gesuchte Erklärung gibt also eine Antwort auf Fragen der Art „Warum existiert in einer Population/Spezies X (heute ausschließlich) die Merkmalsausprägung Y?“, „Warum haben

⁶⁶ Eine andere mögliche Interpretation der Bedingung 2 wäre, b) als eine Erläuterung von a) zu verstehen. In diesem Fall würde Millikan eine probabilistische Theorie der Verursachung vertreten, wie sie von Eells (1991) entwickelt wurde. Nach dessen Theorie ist C dann eine Ursache von F in einem bestimmten Set von Dingen, wenn gilt, dass die Wahrscheinlichkeit F auszuüben unter den Dingen, die C haben, größer ist als unter denjenigen, die nicht C haben, bei gleichbleibenden kausalen Hintergrundbedingungen H. Formal: $P(F|C \& H) > P(F|\neg C \& H)$ (Vgl. Hitchcock 1996, 372). Genau so drückt man aber für gewöhnlich auch Korrelationen aus, d.h. Verursachungen unterscheiden sich nach dieser Theorie nicht von Korrelationen. Mir scheint es jedoch sinnvoll zwischen Korrelationen und Verursachungen zu unterscheiden (Vgl. für Folgendes Sober 2000, 81), denn Korrelationen sind symmetrisch, Verursachungen hingegen nicht. Wenn Rauchen (R) positiv mit Krebs (K) korreliert, dann gilt $P(K|R) > P(K|\neg R)$ aber ebenso muss gelten, dass Krebs (K) positiv mit Rauchen (R) korreliert ist $P(R|K) > P(R|\neg K)$. Wenn Rauchen Krebs verursacht, gilt hingegen nicht zugleich, dass Krebs auch Rauchen verursacht. Korrelationen und Verursachungen sind also prinzipiell auseinanderzuhalten.

⁶⁷ Ein Beispiel für eine Korrelation ohne Verursachung liegt vor, wenn A mit B korreliert, weil beide eine gemeinsame Ursache C haben, aber weder A B verursacht noch umgekehrt B A. Das Fallen des Barometers und das Aufziehen eines Sturms korrelieren miteinander und werden beide durch die momentanen Wetterbedingungen verursacht (Vgl. Sober 2000, 81).

(heute) alle Mitglieder einer Population/Spezies die Merkmalsausprägung Y?“ oder konkreter „Warum haben Eisbären weißes Fell?“

Die Grundidee der Erklärung ist dabei in etwa die Folgende: Ursprünglich existierten in der Population verschiedene Merkmalsausprägungen, neben Y gab es z.B. auch noch die Ausprägung Z. Die Merkmalsausprägungen Y und Z haben sich hinsichtlich ihres Einflusses auf die Überlebensfähigkeit und/oder Reproduktionsfähigkeit ihrer Träger unterschieden. Die Träger der Ausprägungen Y und Z waren ansonsten gleich. D.h. es bestand eine Variation in der Fitness zwischen den Träger von Y und Z. Wenn nun das Verhältnis der prozentualen Anteile der beiden Merkmalsausprägungen in der Population von seinem zeitlichen Verlauf her allein durch die natürliche Selektion bestimmt war, dann ist der Anteil des fitteren Merkmals angestiegen bis er schließlich 100% erreicht hat und ist da geblieben. Was wir heute sehen, so die Erklärung, ist das Resultat des soeben beschriebenen Prozesses, d.h. das Resultat von natürlicher Selektion.

Schauen wir uns nun die Erklärung etwas genauer an.⁶⁸ Dabei gehen wir von einem idealisierten Lebenszyklus aus: Organismen wachsen aus Eiern (Zygoten) zu adulten Formen heran und reproduzieren sich dann, indem sie die nächste Generation an Eiern hervorbringen.

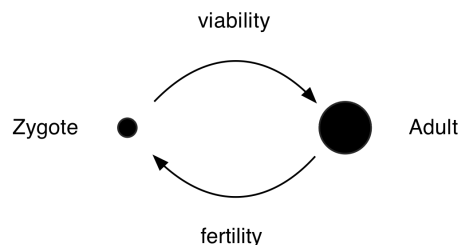


Abbildung 12: Idealisierter Lebenszyklus

Die Wahrscheinlichkeit, dass ein Organismus bis zur Fortpflanzungsfähigkeit überlebt, ist ein Ausdruck für seine Lebensfähigkeit (viability). Die erwartbare Anzahl der Nachkommen, sollte er das fortpflanzungsfähige Alter erreichen, ist hingegen Ausdruck seiner Fruchtbarkeit (fertility). Die natürliche Selektion kann nun entweder bei der Viability oder Fertility oder bei beiden angreifen.

Betrachten wir nun den Fall der Selektion allein aufgrund der Lebensfähigkeit an einem stark vereinfachten Beispiel. Angenommen alle Organismen einer bestimmten Population haben entweder die Merkmalsausprägung Y oder die Merkmalsausprägung Z, sind aber ansonsten gleich. Zum Zeitpunkt t_0 sind die prozentualen Anteile p und q von Y- und Z-trägern im zygoten Zustand gleich verteilt ($p=0,5$ und $q=0,5$). Zudem haben Y und Z einen unterschiedlichen Einfluss auf die Lebensfähigkeit ihrer Träger aber keinen Einfluss auf ihre Fruchtbarkeit, die Wahrscheinlichkeit w_1 , dass Y-Träger das Erwachsenenalter erreichen, liegt bei 0,75, die Wahrscheinlichkeit w_2 für Z-Träger liegt hingegen nur bei 0,5. Wir können nun die Veränderungen der Anteile von Y- und Z-Trägern in der Population innerhalb einer Generation berechnen.

⁶⁸ Für die folgende Darstellung beziehe ich mich maßgeblich auf Sober (2000, 58-60).

	Merkmalsausprägungen	
	Y	Z
Anteile im zygoten Zustand (vor Selektion)	0.5 (p)	0.5 (q)
Fitness	0.75 (w_1)	0.5 (w_2)
Anteile im adulten Zustand (nach Selektion)	0.6 (pw_1/\bar{w})	0.4 (qw_2/\bar{w})

$\bar{w} = pw_1 + qw_2$

Abbildung 13: Veränderung der Anteile nach einer Generation (in Anlehnung an Sober 2000, 59).

Da die Merkmalsausprägungen keinen Einfluss auf die Fruchtbarkeit haben, beträgt der Anteil der Zygoten, die Y-Träger sind, in der nächsten Generation 0,6, ist also um 0,1 angestiegen, während der Anteil der Z-Träger mit 0,4 um 0,1 abgesunken ist. Nach einer weiteren Generation finden wir nur noch ca. 0,3 Z-Träger und schon ca. 0,7 Y-Träger.

	Merkmalsausprägungen	
	Y	Z
Anteile im zygoten Zustand (vor Selektion)	0.6 (p)	0.4 (q)
Fitness	0.75 (w_1)	0.5 (w_2)
Anteile im adulten Zustand (nach Selektion)	0.69 (pw_1/\bar{w})	0.31 (qw_2/\bar{w})

Abbildung 14: Veränderung der Anteile nach der zweiten Generation (in Anlehnung an Sober 2000, 59).

Nach und nach verdrängen also die Y-Träger die Z-Träger bis schließlich nur noch Y-Träger vorhanden sind.

Um Millikans Begriff der Eigenfunktion korrekt zu erfassen, müssen wir Folgendes verstehen: Die Frage, ob ein Effekt positiv, neutral oder negativ ist, und damit eine Eigenfunktion ist oder nicht, hängt wesentlich davon ab, womit er verglichen wird. Etwas vereinfacht kann man sagen, dass der Effekt E1 von Trägern der Eigenschaft p1 positiv ist, wenn der Effekt E2 von Trägern der Eigenschaft p2 geringer zur Entstehung von p2-Trägern beiträgt als E1 zur Entstehung von p1-Trägern, wobei p1-Träger und p2-Träger um dieselben Ressourcen konkurrieren. Ob die Effekte von Trägern reproduktiv etablierter Eigenschaften im konkreten Fall positiv sind oder nicht, ist also in zweifacher Hinsicht relativ. Es hängt nämlich ganz wesentlich davon ab, mit *wem* sie in *welcher Umwelt* tatsächlich um Ressourcen konkurriert haben. D.h. die Beantwortung der Frage, ob die Mitglieder einer bestimmten reproduktiv etablierten Familie eine Eigenfunktion haben, ist zu großen Teilen eine *historische Angelegenheit*. Da uns über die Vergangenheit oft keine genauen Informationen vorliegen, beruht die Zuschreibung einer Eigenfunktion bis zu einem gewissen Grad immer auf Spekulationen und Vermutungen. Es gibt aber auch Beispiele, in denen wir genügend Informationen haben, weil die Evolutionsprozesse relativ kurzfristig abliefen.

Eines dieser Beispiele ist der sogenannte Industriemelanismus, an dem wir uns noch einmal veranschaulichen können, dass die Zuschreibung einer Eigenfunktion davon abhängt, *wer* mit *wem* in *welcher Umwelt* um Ressourcen konkurriert. Kettlewell (1955) konnte bezüglich der Motte (*Biston betularia*) zeigen, dass die Ausbreitung der dunklen Motten (*Biston*

carbonaria) darauf zurückzuführen ist, dass sie auf vom industriellen Russ geschwärzten Birken besser getarnt sind als ihre hellen Verwandten (*Biston betularia*). Dazu führte er u.a. Quasi-Experimente aus, in denen er markierte Motten aussetzte und wieder einfing, wobei er sowohl die Motten (hell/dunkel) als auch die Umgebung (Industriegebiet/ländliche Gegend) variierte (siehe Abbildung 15).

	<i>Biston carbonaria</i>	<i>Biston betularia</i>
Birmingham (städtisch)		
freigelassen	154	64
eingefangen	82	16
% eingefangen	53,2	25
Dorset (ländlich)		
freigelassen	473	496
eingefangen	30	62
% eingefangen	6,3	12,5

Abbildung 15: Anzahl freigelassener und wieder eingefangener Motten in zwei Regionen Englands (in Anlehnung an Maynard-Smith 1992, 40)

Wir können jetzt sagen, dass die dunkle Außenhülle der Motten einen Tarneffekt hat und dass dieser Effekt die Eigenfunktion der dunklen Außenhülle ist.⁶⁹ Dies gilt aber nur relativ zu einer bestimmten Umwelt, wie aus Abbildung 15 klar hervorgeht. Im ländlichen Raum, wo die Birken hell sind, hat hingegen die helle Außenhülle einen Tarneffekt und damit eine Eigenfunktion. Die Zuschreibung der Eigenfunktion ist aber auch relativ zu den Vergleichsobjekten (dunklen und hellen Außenhüllen). Dies wird deutlich, wenn man sich Folgendes überlegt: Eine graue Motte ist auf russgeschwärzter Baumrinde schlechter sichtbar als eine sehr helle Motte, aber besser als eine sehr dunkle Motte. D.h., ob die graue Außenhülle einer Motte eine Eigenfunktion hat, hängt davon ab mit wem, sehr hellen oder sehr dunklen Motten, die grauen Motten historisch konkurrierten.

Wir haben nun ein einigermaßen ausreichendes Verständnis von der evolutionsbiologischen Erklärung entwickelt, auf die sich Millikan in der Definition der Eigenfunktion stützt. Millikans Berufung auf diesen in der Biologie anerkannten Erklärungstyp verleiht ihrer Definition von Eigenfunktion prima facie einen wissenschaftlich seriösen Charakter, welcher dem Funktionsbegriff aufgrund seiner Nähe zur Teleologie oft abgesprochen wurde. Die Diskussion um den wissenschaftstheoretischen Status von teleologischem Denken reicht mindestens bis zu Kants *Kritik der teleologischen Urteilkraft* zurück. Kant vertritt dort die Position, dass die Teleologie als Denken in Endursachen bzw. Zwecken als heuristisches, regulatives Prinzip erlaubt ist, aber nicht die mechanistische Sichtweise der Physik auf die Natur gefährdet, sondern diese bestenfalls ergänzt.⁷⁰

⁶⁹ Die Außenhüllen der Motten betrachten wir dabei als eine reproduktiv etablierte Familie deren reproduktiv etablierte Eigenschaft u.a. eine bestimmte Färbung (z.B. dunkel) ist. Die Färbung hat dann in bestimmten Situationen einen Effekt (Tarnung), der dazu beiträgt, dass weitere dunkle Außenhüllen von Motten entstehen.

⁷⁰ „Um sich also auch nicht der mindesten Anmaßung, als wollte man etwas, was gar nicht in die Physik gehört, nämlich eine übernatürliche Ursache, unter unsere Erkenntnisgründe mischen, verdächtig zu machen, spricht man in der Teleologie zwar von der Natur, **als ob** die Zweckmäßigkeit in ihr absichtlich sei, aber doch zugleich so, dass man der Natur, d.i. der Materie diese Absicht beilegt; wodurch man (weil hierüber kein Missverständnis stattfinden kann, indem von selbst schon keiner einem leblosen Stoffe Absicht in eigentlicher Bedeutung des Wortes beilegen wird) anzeigen will, daß dieses Wort hier nur ein Prinzip der reflektierenden, nicht der bestimmenden Urteilkraft bedeute und also keinen besonderen Grund der Kausalität einführen solle, sondern auch nur zum Gebrauche der Vernunft eine andere Art der Nachforschung ist, als die nach mechanischen Gesetze ist, hinzufüge, um die Unzulänglichkeit der letzteren, selbst zur empirischen Aufsuchung aller besonderen Gesetze zu ergänzen.“ (Immanuel Kant, *Kritik der Urteilkraft*, §68 B 308 (Herv. von A.W.)).

Im 20. Jahrhundert bildet dann die tiefe Beunruhigung über die Teleologie eines der Hauptmotive für die sich entwickelnde philosophische Untersuchung von wissenschaftlichen Erklärungen.

„[...] a deep concern about teleology was a major impetus to the philosophical study to scientific explanation that has developed in the twenty century.“ (Salmon 1989, 111)

Die Beunruhigung wurde zu einem großen Teil von vitalistischen Positionen in der Biologie hervorgerufen und das Ziel von Wissenschaftsphilosophen, wie Braithwaite, Hempel und Nagel, war es Pseudo-Erklärungen, die auf Entelechien oder finale Ursachen Bezug nehmen, als unwissenschaftlich auszuschließen (Vgl. Salmon 1989, 111 u. 26).

Eine solche Bezugnahme finden wir bei Millikan in der Tat nicht. Ihr Begriff von Eigenfunktion weist vielmehr starke Parallelen zum biologischen Begriff der Adaptation auf, den Sober wie folgt definiert:

„Characteristic *c* is an adaption for doing task *t* in a population if and only if members of the population now have *c* because, ancestrally, there was selection for having *c* and *c* conferred a fitness advantage because it performed task *t*.“ (Sober 2000, 85)

Sober weist des Weiteren daraufhin, dass der so definierte Begriff der Adaptation vielleicht der einzige feste Felsen in der bewegten semantischen See („the one firm rock in this shifting semantic sea“, Sober 2000, 87) der biologischen Funktionsbegriffe ist. Funktion verstanden als Adaptation, ist ein Begriff dessen Bedeutung klar genug ist. Jede andere Verwendungsweise des Funktionsbegriffs durch Wissenschaftler oder Philosophen, so Sober, bedarf hingegen einer vorherigen Klärung (Vgl. Sober 2000, 87). In jedem Fall hat der Funktionsbegriff verstanden als Adaptation auch für Sober nichts mit irgendwelchen immateriellen Kräften zu tun, sondern beruht genau wie bei Millikan hauptsächlich auf einer kausal-historischen Beziehung, die durch einen Selektionsprozess geprägt ist.

„Whatever association teleology may have had with vitalism [...] in the past, there is no reason why functional concepts cannot characterize systems that are made of matter and nothing else. The reason the concept of adaption applies to organisms but not to meteors is not that living things contain immaterial ingredients. The difference derives from their very different histories. Selection processes cause some features of objects to be present because they conferred survival and reproductive advantages in the past.“ (Sober 2000, 87)

Insofern der Adaptationsbegriff wissenschaftlich respektabel ist, scheint es auch Millikans Begriff der Eigenfunktion zu sein. Ebenso scheint klar zu sein, wie und wo sich Millikans Begriff der Eigenfunktion in der Biologie verankern lässt.

Diesen Abschnitt abschliessend möchte ich zwei Einwände diskutieren und entkräften, die den Nutzen der evolutionsbiologischen Erklärung für Millikans Definition der Eigenfunktion in Frage stellen.

Der erste Einwand stammt aus Peter McLaughlins *What Functions Explain* (2001). Dort kritisiert McLaughlin die Vertreter von etiologischen Funktionsbegriffen für deren Versuch eine direkte Verbindung zwischen “the selection of X for Y” und “the function of X is Y” etablieren zu wollen (Vgl. McLaughlin 2001, 153-161). Zum Beispiel: Herzen wurden für das Pumpen von Blut (nicht für das Verursachen von Klopfgräuschen) selektiert, also ist die Funktion von Herzen das Pumpen von Blut (nicht das Verursachen von Klopfgräuschen). Das Problem mit dieser Idee ist, dass es den Funktionsträger – das Organ – zur Einheit der Selektion macht. Selektiert werden aber Organismen und nicht Organe. Die etiologische Funktionsauffassung basiert also auf einem Kategorienfehler (Vgl. McLaughlin 2001, 158, 159). Neben dem Begehen eines Kategorienfehlers glaubt McLaughlin zudem, dass die Annahme einer direkten Selektion von Organen (traits) zu seltsamen Funktionszuschreibungen führt. Er bringt hier das Beispiel von Giraffen, die wegen ihrer

langen Beine selektiert wurden und deren lange Hälse das allometrische Beiprodukt der langen Beine sind (Vgl. McLaughlin 2001, 155).

“If traits could be selected directly and if what they were selected for defined their functions, then there would be countless traits whose function was to be morphogenetically linked to useful traits. The long neck of the giraffe would then be selected for its allometric linkage to body size and would have this linkage as its function. If traits are selected for their properties, there is no reason why any property that is causally responsible for the presence of the trait should not be considered its proper function *even if it never did the organism or its progeny any good.*[Hervorh. AW]” (McLaughlin 2001, 157f)

Mit anderen Worten: Wenn die Vertreter etiologischer Funktionsbegriffe konsequent wären, dann müssten sie Funktionen zulassen, deren Ausübung nie einen Anpassungsvorteil hatte. Aber genau dies widerspricht der Kernidee der etiologischen Funktionsauffassung, dass Funktionen etwas mit Selektionvorteilen zu tun haben.

In welchem Ausmass trifft diese Kritik nun auf Millikan zu? Wenn wir die zentralen Elemente von Millikans Definition der Eigenfunktion und von der Erklärung durch Natürliche Selektion⁷¹ gegenüber stellen, dann erhalten wir folgende Übersicht:

elements of an explanation by natural selection		elements of the definition of proper function
population including individuals with trait X and trait Y (giraffe-population with long- legged and short-legged giraffe)	——	Set <i>S</i> (including ancestors of <i>m</i> (things with property <i>C</i>) and things not having <i>C</i>)
trait X (long legs)	——	reproductively established property <i>C</i>
part of the population with trait X (long-legged giraffes)	——	reproductively established family <i>R</i>
adaptive advantage of trait X (to run fast)	——	proper function <i>F</i>
individual with trait X (giraffe with long legs)	——	<i>m</i> (with property <i>C</i>)

Nach dieser Zuordnung scheint Millikan von einer Selektion von Organen auszugehen. Millikan ist aber sicherlich nicht zu dieser Annahme verpflichtet. Alles, was Millikan voraussetzt, ist, dass es Erklärungen gibt, die die Existenz bestimmter Organe über deren Korrelation mit bestimmten Effekten erklären.

„Historical statistical correlations between organ or behavior structures and the performances or reliability of the performances of certain useful functions thus figure in legitimate explanations of why the modern Normal traits of any species have continued to proliferate thus culminating in the existence of current exemplars of these traits.“ (Millikan 1984, 27)

D.h. aber, dass die Erklärung, die die Existenz der langbeinigen Giraffen erklärt *zugleich* auch die Existenz der langen Giraffenbeine erklärt. Wichtig ist hierbei bloss, dass die Erklärung den Effekt der langen Beinen berücksichtigt, aber dies impliziert nicht notwendigerweise eine Selektion von Organen (parts). Der Effekt (schnelleres Laufen) ist die Eigenfunktion der langen Beine. Genau diese Auffassung scheint McLaughlin aber selbst zu akzeptieren:

⁷¹ Zur Veranschaulichung wurde in der Übersicht angenommen, dass sich in einer Giraffenpopulation mit ursprünglich kurz- und langbeinigen Giraffen allnählich die langbeinigen Giraffen durchgesetzt haben, weil sie schneller vor Fressfeinden fliehen konnten.

“At best, we could say that our attribution of the function Y to X is based on our causal analysis of X 's effects, which indicates that doing Y has been X 's contribution to an organism's reproductive success.” (McLaughlin 2001, 157)

Mehr als das benötigt Millikan aber nicht. Daraus folgt, dass sich McLaughlins Kritik an Millikans Definition von Eigenfunktion vermeiden lässt. Wir wollen nun noch kurz darlegen, dass auch McLaughlins zweiter Einwand, das Beispiel mit den langbeinigen und langhalsigen Giraffen, bei Millikan zu keiner seltsamen Funktionszuschreibung führt.

Nach Millikan sind Beine und Hälsen Mitglieder von reproduktiv etablierten Familien höherer Ordnung. Wir werden darauf noch ausführlich unter Abschnitt 2.1.1.3.5 eingehen. Hier reicht es, wenn wir wissen, dass Mitglieder von reproduktiv etablierten Familien höherer Ordnung (HOREF A) die Produkte von Mitgliedern reproduktiv etablierter Familien erster Ordnung (FOREF X) sind, deren Eigenfunktion es ist, die Ersteren zu produzieren (siehe Abbildung 9).

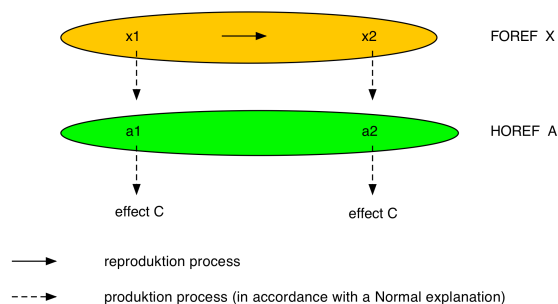


Abbildung 16: Entstehung von Mitgliedern einer reproduktiv etablierten Familie höherer Ordnung (HOREF A)

Der Einfachheit halber nehmen wir nun an, dass die Hälsen (HOREF N) und die Beine (HOREF L) die Produkte der FOREFs A bzw. B sind und dass deren Bildung durch ein Protein P (HOREF P), das das Produkt der FOREF C ist, beeinflusst wird. Das Protein P bewirkt, dass die Hälsen und Beine kurz sind (die Eigenschaft S haben) und somit die Effekte E1 und E2 haben (siehe Abbildung 17).

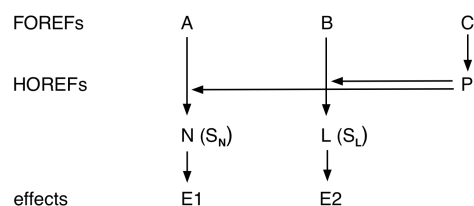


Abbildung 17: Ursachen und Wirkungen kurzer Hälsen und Beine (fiktives Beispiel)

Nehmen wir nun an, dass eine Mutation in der Giraffenpopulation auftritt: Die FOREF C ändert sich zu FOREF D, deren Produkt das Protein P* ist. P* bewirkt nun die Bildung von langen Hälsen und Beinen, und diese haben die Effekte E3 und E4 (siehe Abbildung 18).

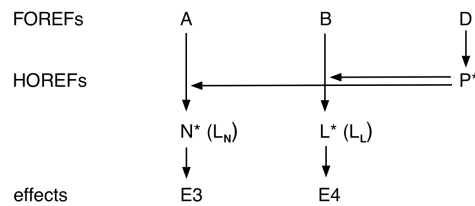


Abbildung 18: Ursachen und Wirkungen langer Hälse und Beine (fiktives Beispiel)

Nach einer gewissen Zeit setzen sich die langhalsigen und langbeinigen Giraffen durch, weil deren lange Beine (ex hypothesis) einen Selektionsvorteil darstellen. D.h. der Effekt E4 (schnelleres Laufen) verursacht statistisch einen höheren Reproduktionserfolg als der Effekt E2. Die Effekte E1 und E3 sind hingegen bezüglich ihres Einflusses auf den Reproduktionserfolg gleich.

Nach Millikan haben dann die langen Beine der Giraffen eine Eigenfunktion, nämlich, ein schnelleres Laufen zu ermöglichen, weil alle drei Bedingungen ihrer Definition erfüllt sind.

Was ist nun aber mit den langen Hälsen? Könnte deren Funktion der Effekt E3 sein? Nein, nicht in Übereinstimmung mit Millikans Definition, denn die dritte Bedingung der Definition ist nicht erfüllt. Es kann keine Erklärung der relevanten Art geben, weil die Effekte E1 und E3 keinen Unterschied bezüglich des Reproduktionserfolgs der Giraffen machen. Könnte der Effekt E4 die Funktion der langen Hälse sein? Nein, denn dann wäre die zweite Bedingung der Definition nicht erfüllt. Es gibt keine direkte kausale Verbindung zwischen der Eigenschaft lange Hälse zu haben (L_N) und dem Effekt E4. Könnte die allometrische Verbindung (allometric linkage) zu den langen Beinen die Funktion der langen Hälse sein, wie dies McLaughlin unterstellt? Auch dies funktioniert nicht. Dafür müsste die allometrische Verbindung (allometric linkage) ein Effekt der langen Hälse sein, was kaum Sinn ergibt. Wenn die allometrische Verbindung (allometric linkage) zu langen Beinen als Eigenschaft der langen Hälse ein Selektionsvorteil wäre, dann hätten die langen Hälse eine Eigenfunktion. Aber diese Eigenfunktion bestünde in irgendeinem Effekt und nicht in der allometrischen Verbindung (allometric linkage) selbst, wie die McLaughlin unterstellt.

D.h., Millikans Definition von Eigenfunktion führt zu dem korrekten Resultat, dass die langen Hälse keine Eigenfunktion haben. Damit ist McLaughlins Kritik insgesamt entkräftet. Sein erster Einwand lässt sich vermeiden und der zweite trifft nicht zu: Millikan ist weder zur Annahme einer Selektion von Organ (parts) verpflichtet noch führt ihre Definition zu seltsamen Funktionszuschreibungen.

Eine andere Schwierigkeit bei der Übertragung der oben dargestellten evolutionsbiologischen Erklärung in den Kontext von Millikans Definition der Proper Function liegt in einer Verschiebung des Explanandums. Die evolutionsbiologische Erklärung, so hatten wir gesagt, kann erklären, warum (heute) alle Mitglieder einer Population/Spezies die Merkmalsausprägung Y haben, bzw., warum in einer Population/Spezies X (heute ausschließlich) die Merkmalsausprägung Y existiert, konkret z.B., warum Eisbären weißes Fell haben.

In Millikans Definition der Proper Function soll die Erklärung aber etwas anderes leisten. Sie soll erklären warum ein *einzelnes Individuum* mit der Merkmalsausprägung Y existiert, z.B., warum der Eisbär Knut ein weißes Fell hat.

Es findet also ein Wechsel des Explanandums statt, von:

- a) der Existenz einer Merkmalsausprägung bei allen Mitgliedern einer Population
- zu:

- b) der Existenz einer Merkmalsausprägung bei einem Individuum bzw. die Existenz einer Merkmalsausprägung als Token⁷².

Dieser Wechsel scheint insofern problematisch als Elliot Sober in *The Nature of Selection* explizit dafür argumentiert, dass die evolutionsbiologische Erklärung⁷³ nur a) aber nicht b) erklären kann (Vgl. Sober 1984a, 147-155).

„The frequency of traits in a population can be explained by natural selection, even through the possession of those traits by individuals in the population cannot.“ (Sober 1984a, 152)

Um seinen Standpunkt klar zu machen unterscheidet Sober zwischen zwei Arten von Erklärungen, die eine nennt er *developmental* und die andere *selectional* (oder *variational*). Den Unterschied zwischen beiden erläutert er an einem Beispiel. Angenommen in einem Klassenraum befinden sich vier Schüler (Sam, Aaron, Martin und Alexander), die alle auf Drittklässlerniveau lesen können. Diesen Umstand können wir auf zweierlei Weise erklären:

- 1) Developmental. Wir erklären den Umstand, dass alle Schüler in diesem Raum auf Drittklässlerniveau lesen, dadurch, dass wir beschreiben wie relevante Erfahrungen und Bedingungen bei jedem dieser Schüler einzeln dazu geführt haben, dass er die Fähigkeit entwickelt/erworben hat auf Drittklässlerniveau zu lesen.
- 2) Selectional. Wir erklären den Umstand, dass alle Schüler in diesem Raum auf Drittklässlerniveau lesen, dadurch, dass es eine Eintrittsbedingung für diesen Raum ist auf Drittklässlerniveau lesen zu können (In einem Vorzimmer des Raumes machen alle Schüler einen Lesetest und es dürfen nur diejenigen eintreten, die auf Drittklässlerniveau lesen).

Beim zweiten Erklärungstyp (*selectional*), der unserer evolutionsbiologischen Erklärung entsprechen soll, wird anders als beim ersten Erklärungstyp (*developmental*) nicht erklärt, warum ein einzelner Schüler seine Lesefähigkeit hat. Dasselbe soll auch für biologische Beispiele gelten.

„In the schoolroom example, selection does not explain why any child can read at the third grade level. The fact that an individual now in the room can do this is accounted for by the stability assumption: That individuals could read at the third grade level at some earlier time, and the trait persisted. Selection is not what does the explanatory work here. The same holds true for biological examples. Natural selection does not explain why I have an opposable thumb (rather than lack one). This fact falls under the purview of the mechanisms of inheritance [...] But phenotypic and genotypic properties of individuals – properties of morphology, physiology, and behavior – fall outside natural selection’s domain.“ (Sober 1984a, 152)

Sobers Behauptung, dass die Natürliche Selektion die Häufigkeiten eines Merkmals in einer Population erklären kann, aber nicht das Haben eines Merkmals von einzelnen Individuen in der Population, hat Karen Neander in *What Does Natural Selection Explain? Correction to Sober* (1988) widersprochen. Neander weist auf zwei Doppeldeutigkeiten in Sobers Behauptung hin. Erstens kann man Natürliche Selektion eng und weit verstehen. Eng verstanden ist damit bloße Selektion gemeint, d.h. differentielles Überleben und Reproduzieren von Individuen in einer Population. Weit verstanden ist damit bloße Selektion plus Vererbung und die Entstehung von Variation durch Mutation und Rekombination gemeint. Zweitens, kann „erklären“, „vollständig erklären“ oder „teilweise erklären“ heißen. Neander versteht Sober so, dass er behauptet die Natürliche Selektion (eng verstanden) erklärt (als teilweises Erklären verstanden) nicht die Merkmale eines Individuums.

⁷² Warum hat Knut der Eisbär ein weißes Fell bzw. warum existiert Knuts weißes Fell?

⁷³ Statt von einer evolutionsbiologischen Erklärung spricht Sober von einer Erklärung durch Natürliche Selektion.

„So I take Sober to be saying that mere selection is no part of explanation of the phenotypic and genotypic traits of individuals.“ (Neander 1988, 424)

Sie stimmt Sober auch darin zu, dass das Merkmal eines Individuums (z.B. das Haben eines opponierbaren Daumens) zunächst durch die Mechanismen der Vererbung, d.h. „genetic coding, meiosis and fertilization“ (Neander 1988, 425), erklärt werden kann. Wir können aber, so Neander, weitergehen und fragen, woher die Eltern des Individuums die Eigenschaft haben und deren Eltern und deren Eltern usw. Die Antwort auf diese rekursive Frage geht aber nicht ins Unendliche, vielmehr finden wir auf der Linie der Vorfahren irgendwann Individuen, die so etwas wie eine Vorstufe eines opponierbaren Daumens haben. Verfolgen wir die Ahnenreihe noch weiter in die Vergangenheit, dann finden wir irgendwann Individuen, die so etwas wie die Vorstufe der Vorstufe eines opponierbaren Daumens haben usw.

„The answer to the recursive question is not that their parents inherited the genotype from theirs, and they form theirs, and so on, ad infinitum. This answer *would* remain within the scope of the mechanisms of inheritance, but it is *false*. The correct explanation does not go on in this way ad infinitum. Somewhere back along the line [...] there were occasional mutations of significant genetic recombination.“ (Neander 1988, 425)

Rückblickend führen die Veränderungen in den Vorstufen zur Ausbildung von opponierbaren Daumen (siehe Abbildung 19). Und gerade weil die Veränderungen zufällig sind, muss die Abfolge der Vorstufen durch die Natürliche Selektion erklärt werden. In diesem Sinne spielt die Natürliche Selektion auch bei der Erklärung der Eigenschaften eines Individuums eine Rolle.

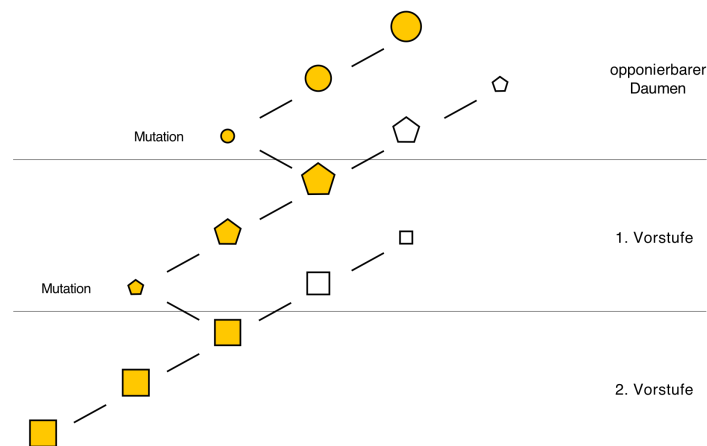


Abbildung 19: Stark schematisierte Darstellung der Entwicklung zum opponierbaren Daumen (die Grösse der Symbole reflektiert die Anteile in der Population).

Neander fasst ihr Argument wie folgt zusammen:

„My point is that the explanation of a genotypic or phenotypic property of an individual involves explaining the frequency of genotype and phenotype in a population. The population in question is the ancestors of the individual whose inherited characteristic is being explained. To stop the explanatory description of its inheritance at its parents is arbitrary. A most pertinent fact about its extended ancestry is the high frequency among them of certain kind of phenotype and genotype. That is, in the case of our

opposable thumb, there is a great predominance of individuals who inherited the more opposable thumb of the varieties available at the time, or who developed directly from sex cells that mutated or genetically recombined in this direction. This latter kind of fact is one that Sober himself claims it is the role of natural selection to explain. It is explained by selective force in favor of individuals with more opposable thumbs. The others being selected against, bred less successfully in terms of fitness, and so our ancestry is not traceable to them (where their line dies out completely) or is traced less frequently to them. Thus a more complete explanation of an individual's traits goes beyond the mechanisms of inheritance to those of natural selection.“ (Neander, 426)

Sollte Neander mit dieser Argumentation Recht haben, was m.E. der Fall ist⁷⁴, dann ist Millikans Wechsel des Explanandums unproblematisch.

⁷⁴ Man könnte Neander hier so verstehen, dass sie Sober darauf hinweist, dass er ohne erkennbare Begründung nur proximate Ursachen, aber keine ultimativen Ursachen bei der Erklärung berücksichtigt. Zur Unterscheidung von proximate und ultimativen Ursachen siehe Mayr (1961). Glock (####) unterscheidet noch genauer in proximate, distale und ultimate Ursachen.

2.1.1.3.4 Normal Explanation und Normal Condition

In diesem Abschnitt sollen Millikans Begriffe „Normal Explanation“ und „Normal Condition“ eingeführt und erläutert werden. Beginnen wir mit der Normalen Erklärung⁷⁵.

Nach Millikan muss es zu jeder Eigenfunktion eine *Normale Erklärung* geben. Eine solche Erklärung erklärt, *wie* eine bestimmte reproduktiv etablierte Familie in der Vergangenheit eine bestimmte Eigenfunktion ausgeführt hat.

„A Normal Explanation is an explanation of how a particular reproductive established family has historically performed a particular proper function.“ (Millikan 1984, 33)

Dies stellt sicher, dass Eigenfunktionen keine Effekte sind, die bloss zufällig wiederholt aufgetreten sind.

Nach Millikan weisen solche Erklärungen Grade an Umfang und Genauigkeit bzw. Nähe (proximity) auf, wobei die am engsten begrenzte (most proximate)⁷⁶ und damit am wenigsten genaue Normale Erklärung für eine reproduktiv etablierte Familie *R* und eine Eigenfunktion *F* mindestens Folgendes umfassen muss:

1. Angabe der Eigenschaften der Mitglieder der reproduktiv etablierte Familie *R*
2. Angabe der Bedingungen unter denen die Mitglieder der reproduktiv etablierte Familie *R* in der Vergangenheit tatsächlich die Eigenfunktion *F* ausgeführt haben
3. Angabe eines oder mehrerer Naturgesetze
4. Ableitung bzw. lückenloses detailliertes Aufzeigen, wie 1., 2. und 3. (this set up) zur Ausübung der Eigenfunktion *F* führt (Vgl. Millikan 1984, 33)⁷⁷.

Millikan veranschaulicht die verschiedenen Genauigkeitsgrade am Beispiel Normaler Erklärungen der Funktion des Herzens:

„For example, the most proximate Normal explanation of how the human heart circulates blood must tell something of how the heart is made (its Normal properties or structure), how it works inside, and note such conditions as the regularity of electrical impulses sent to the heart, the oxygen supply sent to the heart, and the presence of a closed circuit of blood vessels emanating from and returning to the heart which carry the blood to and from the proper places in the heart. A slightly less proximate Normal explanation will tell where the electrical impulses have usually come from, then how they in turn are usually produced. The heart of a person who must wear a pacemaker

⁷⁵ Millikan verwendet hier ein Grosses N um deutlich zu machen, dass normal nichts mit statistischem Durchschnitt zu tun hat.

⁷⁶ Aufgrund der in der Biologie weitverbreiteten Unterscheidung zwischen proximaler und ultimer Erklärung (Tinbergen 1963) ist dies eine eher unglückliche terminologische Wahl. Diesen Hinweis verdanke ich Hanjo Glock.

⁷⁷ Etwas unklar ist, was für einem Erklärungstyp Normale Erklärungen entsprechen sollen. Millikans Redeweise, dass die Ausübung der Funktion *F* aus den Angaben (1-3) *abgeleitet* werden soll und die Forderung, Naturgesetze anzugeben, verweisen auf das Modell der DN-Erklärung von Hempel. Aufgrund einiger Schwierigkeiten gilt dieses Modell der wissenschaftlichen Erklärung heute unter Wissenschaftsphilosophen als ein Relikt des Logischen Positivismus (Vgl. Craver 2009, 35, Salmon 1989). Millikan ist aber, wie bereits erwähnt (Abschnitt 2.1.1.3.1), nicht zwingend auf Erklärungen dieses Typs angewiesen, vielmehr genügt es, wenn es überhaupt eine wissenschaftlich akzeptable Erklärung für das jeweilige Explanandum gibt. Eine solche Erklärung könnte auch eine mechanistische Erklärung im Sinne von Salmon 1989, Bechtel/Richardson 1993, Machamer et. al. 2000, Craver 2009 sein.

does not pump blood in accordance with a Normal explanation except on the most proximate level. A less proximate Normal explanation might also tell where the oxygen usually comes from.“ (Millikan 1984, 33)

Die Eigenschaften, die unter 1. aufgeführt werden müssen, nennt Millikan *Normal functioning Properties* und die Bedingungen aus 2. bezeichnet sie als *Normal conditions*.

Normale Bedingungen sind diejenigen Bedingungen, die in der Normalen Erklärung einer Eigenfunktion *erwähnt werden müssen*.

„The conditions that must be mentioned [in such an explanation, AW] are „Normal condition“ for the proper performance of members of R.“ (Millikan 1984, 33)

Der Begriff der Normalen Bedingung ist also immer relativ zu einer bestimmten Eigenfunktion zu verstehen und sein Umfang/seine Extension hängt zudem davon ab, welchen Grad an Detailliertheit die Normale Erklärung der Eigenfunktion hat. Ich beschränke meine Erläuterungen hier auf die am wenigsten detaillierte Normale Erklärung (most proximate Normal explanation).⁷⁸

Wir können uns nun an unserem Beispiel des Industriemelanismus (siehe S. 61f) klar machen, welche Bedingungen Normale Bedingungen sind und welche nicht. Angenommen wir wären in der Lage gewesen alle Bedingungen zu erfassen, in denen die dunklen Außenhüllen von Motten in der Vergangenheit einen Tarneffekt gehabt haben, dann stellt sich die Frage, welche dieser Bedingungen Normale Bedingungen sind. Wir könnten hier zunächst die relevanten von den irrelevanten Bedingungen unterscheiden, auch wenn diese Unterscheidung nicht trennscharf ist. Vielleicht war z.B. bei jeder Ausübung der Tarnfunktion immer eine bestimmte Mäuseart im Umkreis von 50m um die betreffende Motte. Selbst wenn dies kein Zufall ist⁷⁹, so ist es doch bis auf weiteres irrelevant, um das Eintreten des Tarneffektes zu erklären. Ebenso irrelevant wäre es, wenn alle Tarneffekte am Nachmittag und nie am Vormittag stattgefunden hätten. Wahrscheinlich herrschte in allen Situationen der Ausübung des Tarneffektes aber auch ein bestimmtes Temperaturspektrum vor, eine gewisse Strahlungsintensität, eine gewisse Gravitation, ein bestimmter Sauerstoffgehalt der Luft etc. Alle diese Bedingungen sind nicht irrelevant für die Ausübung der Tarnfunktion, denn, ohne dass diese Bedingungen vorliegen, können weder die Motten noch deren Fressfeinde überhaupt existieren. Innerhalb der relevanten Bedingungen können wir nun zwischen den spezifischen und den unspezifischen Bedingungen unterscheiden. Unspezifische Bedingungen sind diejenigen, die generell vorliegen müssen, um den Motten die Existenz zu ermöglichen. Die oben genannten Bedingungen sind also unspezifische Bedingungen. Spezifische Bedingungen hingegen sind diejenigen, die u.a. die Ausübungssituation ausmachen. Es sind diejenigen Bedingungen, unter denen sich der Vorteil der Eigenschaft dunkel zu sein gegenüber der Eigenschaft hell zu sein zeigt. Zu diesen Bedingungen gehören in unserem Beispiel mindestens die Bedingungen „Motte sitzt auf dunklem Untergrund“ und „Motte befindet sich im Blickfeld eines Fressfeindes“. Diese Bedingungen sind Normale Bedingungen, denn sie müssen selbst bei der größten Erklärung, wie der Tarneffekt zustande kommt erwähnt werden. Eventuell müssen diese Bedingungen noch ergänzt werden. Vielleicht dürfen die Sichtverhältnisse für den Fressfeind nicht optimal sein, weil er sonst trotz der dunklen Färbung in der Lage ist die Umrisse der Motte zu erkennen oder der Fressfeind muss hungrig sein etc. Dass wir die Normalen Bedingungen fast nie vollständig erfassen können, heißt aber nicht, dass wir nicht zumindest einige dieser Bedingungen kennen. Liegt eine der uns bekannten Normalen Bedingungen nicht vor, dann

⁷⁸ Bedingungen, die in dieser Erklärung erwähnt werden müssen, müssen auch in detailreicheren Erklärungen erwähnt werden. Wir erhalten durch diese Beschränkung also so etwas wie die minimalen Normalen Bedingungen.

⁷⁹ Es könnte ökologische Zusammenhänge geben, die erklären, warum sich die Mäuse in der Nähe der Motten aufhalten.

befindet sich der betreffende Funktionsträger in abNormalen Bedingungen. Dies wird später noch von Bedeutung sein (Vgl. Abschnitt 3.2.9).

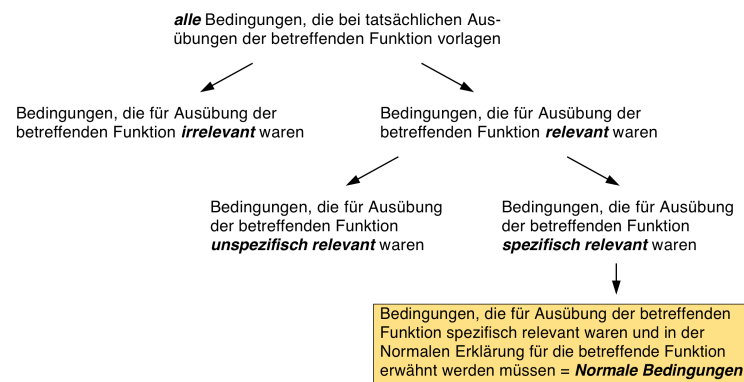


Abbildung 20: Einordnung der Normalen Bedingungen unter allen Bedingungen, die bei tatsächlichen Ausübungen der Funktion immer vorlagen

Wir haben jetzt ein hinreichend genaues Verständnis davon, was eine Normale Erklärung ist und was Normale Bedingungen sind.

2.1.1.3.5 Higher order reproductively established families

In diesem Abschnitt soll Millikans Definition von einer higher order reproductively established family (HOREF) eingeführt und erläutert werden.

Millikan gibt drei unterschiedliche Definitionen zu den reproduktiv etablierten Familien höherer Ordnung. Die letzte der drei Definitionen soll aber eher eine Lockerung der vorangegangenen Definitionen darstellen um die Möglichkeit von missgebildeten REF-Mitgliedern einzuräumen, daher betrachten wir hier nur die ersten beiden Definitionen. Der Möglichkeit von missgebildeten REF-Mitgliedern wenden wir uns später unter Abschnitt 2.2.3.1 noch gesondert zu, denn auf dieser Möglichkeit basiert einer der beiden Ansätze von Millikan zur Lösung des Problems der Missrepräsentation.

Die beiden relevanten Definitionen lauten:

„(1) Any set of similar items produced by members of the same reproductively established family, when it is a direct proper function of the family to produce such items and these are all produced in accordance with Normal explanations form a *higher-order reproductively established family*.“ (Millikan 1984, 24)

„(2) Any set of similar items produced by the same device, when it was one of the proper functions of this device to make later items *match* earlier items, and these items are alike in accordance with a Normal explanation for the performance of this function, form a *higher-order reproductively established family*.“ (Millikan 1984, 24)

Mit (1) sollen nach Millikan vor allem Organe, wie Herz oder Leber, aber auch starre Verhaltensäußerungen, wie Balzrituale oder Fellpflege, als Mitglieder von HOREFs ausgewiesen werden. (2) soll hingegen auch Äußerungen individuell erlernten Verhaltens als Mitglieder von HOREFs erfassen (Vgl. Millikan 1984, 25).

Wir hatten bereits bezüglich der Kriterien für die Mitgliedschaft in einer FOREF festgestellt, dass es sich hierbei um eine Kombination aus dem Vorhandensein einer bestimmten kausal-historischen Verbindung, die zudem ein gewisses Maß an Regelmäßigkeit aufweisen muss (sonst gibt es keine Erklärung), und dem Vorhandensein einer Gleichheit bzw. Ähnlichkeit hinsichtlich bestimmter Eigenschaften handelt (siehe Abschnitt 2.1.1.3.2). Dasselbe gilt auch für die reproduktiv etablierten Familien höherer Ordnung. Allerdings ergibt sich bei den HOREF die kausale Verbindung und die Ähnlichkeit nicht direkt, sondern nur indirekt durch einen wissenschaftlich erklärbaren Reproduktionsmechanismus. Der Reproduktionsmechanismus erklärt die Ähnlichkeit der FOREF-Mitglieder, die Ähnlichkeit der HOREF-Mitglieder als deren Produkte wird hingegen durch die Normale Erklärung garantiert. D.h., die kausal-historische Verbindung zwischen den HOREF-Mitgliedern geht immer über die dazugehörigen FOREF-Mitglieder. Die Mitglieder einer HOREF können aber auch die Produkte von Mitgliedern einer anderen HOREF sein, es gibt also REFs verschieden hoher Ordnung und nicht nur REFs erster und zweiter Ordnung⁸⁰. Die Zusammenhänge zwischen FOREFs und HOREFs vor allem im Sinne der ersten Definition können wir uns anhand eines Schemas verdeutlichen:

⁸⁰ Letzteres scheint Wild zumindest indirekt durch seine Ausdrucksweise zu implizieren (Wild 2010, 52).

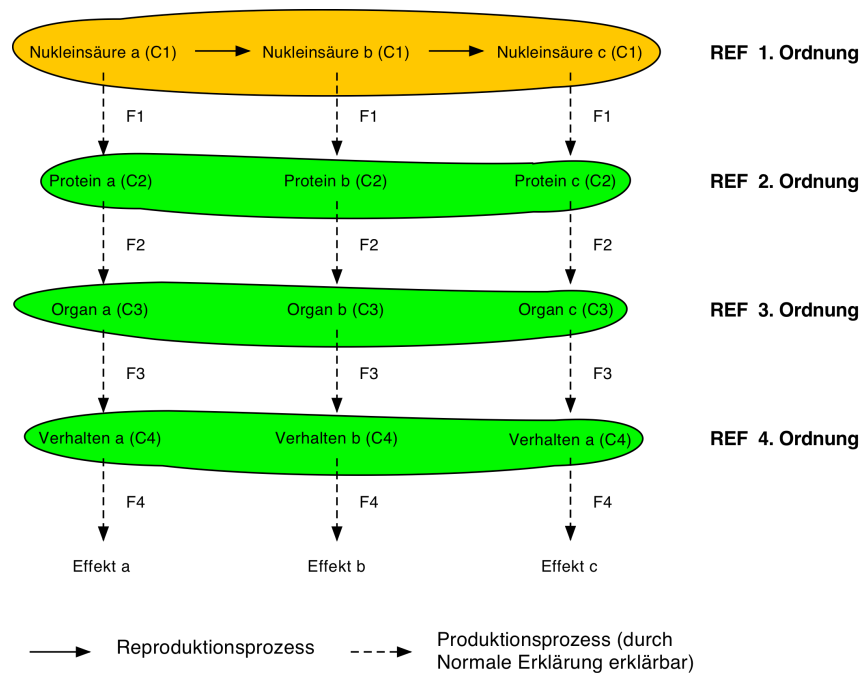


Abbildung 21: Schematische Darstellung des Zusammenhangs zwischen FOREFs und HOREFs

Erläuterung des Schemas:

Die Nukleinsäuren a, b und c bilden eine FOREF relativ zu reproduktiv etablierten Eigenschaft C1, dem Codieren für das Protein X. Ein Effekt der Eigenschaft C1 ist dann die Synthese des Proteins X. Dieser Effekt ist zugleich die Eigenfunktion F1 der Nukleinsäuren a, b und c, sofern der Effekt eine Selektion gegenüber Nukleinsäuren ohne C1 bewirkt hat. Zudem kann das Zustandekommen des Effekts durch die Nukleinsäuren mittels einer Normalen Erklärung erklärt werden⁸¹. Die Proteine a, b und c sind also die Produkte von Mitgliedern einer REF erster Ordnung deren Funktion es ist solche Produkte hervorzubringen und die Produktion ist einer Normalen Erklärung zugänglich. Somit bilden die Proteine a, b und c eine REF zweiter Ordnung. Der Synthesemechanismus, den die Normale Erklärung anführt, besitzt eine gewisse Regelmäßigkeit, so dass die Proteine a, b und c nahezu gleiche Eigenschaften haben. Unter diesen Eigenschaften mag eine Eigenschaft C2 sein, die einen fitnesserhöhenden Effekt von C2-Trägern gegenüber Nicht-C2-Trägern hat. Wir könnten hierbei z.B. an eine katalytische Wirkung der Proteine vom Typ X für den Aufbau der das Protein codierenden Nukleinsäuren denken. Der Effekt von C2 könnte aber auch im Aufbau von bestimmten Gewebestrukturen oder ganzen Organen bestehen. Hierfür wäre natürlich die Synthese von mehreren Proteinen des Typs X (und einigem mehr) notwendig.

Aus den bis hier hin anhand des Schemas entwickelten Vorstellungen ergeben sich zwei Schwierigkeiten, die einer Erläuterung bedürfen. Die erste Schwierigkeit besteht in einer gewissen Arbitrarität beim Bestimmen der Eigenfunktion der Funktionsträger. Eigenfunktionen, so hatten wir gesagt, sind Effekte, die dem Funktionsträger einen Selektionsvorteil verschaffen und wir hatten bezüglich des Schemas angenommen, dass die Eigenfunktion der Nukleinsäuren in der Synthese eines Proteins besteht und die Eigenfunktion der Proteine im Aufbau eines Organs oder der Katalyse der sie codierenden Nukleinsäure. Da Kausalrelationen transitiv sind, gilt, dass die Ursache eines Effektes, der einen weiteren Effekt verursacht, auch die Ursache des zweiten Effektes ist. Wir könnten also

⁸¹ Die Normale Erklärung bestünde in diesem Fall in der Beschreibung der sogenannten Proteinbiosynthese, ein Prozess der mittlerweile sehr gut erforscht ist.

bezüglich der Nukleinsäuren genauso gut behaupten, dass deren Eigenfunktion darin besteht weitere Nukleinsäuren ihres Typs zu katalysieren. Diese Interpretation scheint umso plausibler, als es letztlich dieser Effekt ist, der für den selektiven Vorteil verantwortlich ist. Die Synthese des Proteins allein ist hingegen dafür nicht ausreichend. Für den Fall, die Proteine bewirken den Aufbau eines Organs müssten wir ebenso annehmen, dass das Organ direkt einen selektionsvorteilverschaffenden Effekt hervorruft oder aber der Effekt des Effektes des Organs (d.i. das Verhalten). Dieser Logik folgend können wir sämtlichen REF-Mitgliedern im Schema (Nukleinsäuren, Proteine, Organe, Verhalten) die Eigenfunktion F4 zuschreiben. Millikan ist sich dieser Problematik bewusst und bietet eine Lösung an.

Nach Millikan können REF-Mitglieder auch mehrere Eigenfunktionen haben. Je nachdem, ob diese Eigenfunktionen unabhängig voneinander ausgeübt werden oder nicht, spricht Millikan von alternativen (unabhängigen) oder simultanen (abhängigen) Eigenfunktionen. Die Federn von Wasservögeln (device) haben nach Millikan alternative Eigenfunktionen, sie halten den Organismus warm (A), dienen der fliegenden Fortbewegung (B) und spielen eine Rolle bei Paarungsritualen (C) (siehe Abbildung 22). Die Versorgung anderer Organe mit Sauerstoff und die gleichzeitige Versorgung mit Nährstoffen sind hingegen simultane Eigenfunktionen des Herzens.

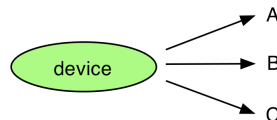


Abbildung 22: Organ (device) mit drei alternativen Eigenfunktionen (Beispiel: Gefieder von Wasservögeln)

Zudem kann ein device mehrere Eigenfunktionen haben, weil sie eine Kette von Effekten bilden. Diese Eigenfunktionen nennt Millikan serielle Eigenfunktionen (siehe Abbildung 23)⁸².

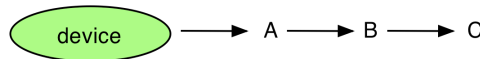


Abbildung 23: Organ (device) mit drei seriellen Eigenfunktionen (Beispiel: menschliches Herz)

In diesen Fällen von mehreren Eigenfunktionen kann man nun fragen: Was ist aber *die* Eigenfunktion der Vorrichtung?

Nach Millikan hängt die Antwort auf die Frage, was die Eigenfunktion einer Vorrichtung ist, vom Verkettungsmuster der Eigenfunktionen ab, die die betreffende Vorrichtung (device) hat. Schauen wir uns dazu zwei Beispiele an, in denen Kombinationen von seriellen und alternativen bzw. simultanen Eigenfunktionen vorkommen.

1) Das Herz (device) pumpt Blut (A) und transportiert dadurch Sauerstoff (B) und Nährstoffe (C) durch den Körper. Dies erhält die Leber (D), das Gehirn (E) und die Lunge (F) am Leben. Die ihrerseits Eigenfunktionen haben (nicht dargestellt).

⁸² Millikan gibt hier folgendes Beispiel: "For example, the human heart pumps blood; this causes oxygen to circulate through the blood vessels; this causes oxygen to reach the brain, thus helping brain cells to live, thus helping the brain to think properly, thus causing appropriate human actions and utterances, etc. Call functions that are related to one another in this way 'serial functions.'" (Millikan 1984, 37)

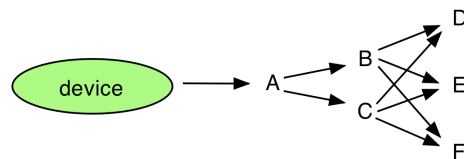


Abbildung 24: Verkettungsmuster der Eigenfunktionen des Herzens

2) Ein fiktives Beispiel von Millikan (Vgl. Millikan 1984, 36f): Spezielle kleine Zellen im Hypothalamus (device) bewirken einen erhöhten Blutfluss durch die Hautkapillaren (A), die Absonderung von Schweiß durch die Schweißdrüsen (B) und eine Verminderung der Aktivität (C). Die Effekte A, B und C bewirken gemeinsam, dass die Körpertemperatur von 98,6 ° Fahrenheit (D) wiederhergestellt wird.

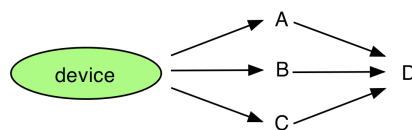


Abbildung 25: Verkettungsmuster der Eigenfunktionen spezieller Zellen (fiktives Beispiel)

In Fällen, die ein ähnliches Muster wie 1) zeigen, wo es also mehrere Endpunkte (D, E, F) gibt, meint Millikan, dass wir dazu tendieren den ersten Effekt als die Eigenfunktion des devices aufzufassen⁸³. Die Eigenfunktion des Herzens wäre es also Blut zu pumpen (A).

„Something like explanatory proximity seems to be involved here. In the order of explanation, the heart pumps blood first and only therefore can do other things.“
(Millikan 1984, 36)

In Fällen, die hingegen dem Beispiel 2) ähneln, wo es nur einen Endpunkt gibt, ist es das letzte Glied, dem nach Millikan unter den Eigenfunktionen eines devices eine besondere Rolle zukommt. Diese letzte Funktion bezeichnet Millikan als Focused Proper Function.

„The last member of such a series of functions I will call the focused function of such a device.“ (Millikan 1984, 36)

Die Fokussierte Eigenfunktion der kleinen Zellen im Hypothalamus wäre also die Konstanthaltung der Körpertemperatur bei 98,6° Fahrenheit⁸⁴.

Letztlich bietet die hier angegebene Erläuterung aber nur eine grobe Orientierung um eventuell die fokussierte Eigenfunktion eines devices zu bestimmen, denn nach Millikan kann ein device auch mehrere fokussierte Eigenfunktionen haben. In einem solchen Fall wäre unser bisheriges Kriterium (ein Endpunkt vs. mehrere Endpunkte) ausgehebelt.

„I do not, however, wish to make it actually *contradictory* that a device *might* have, say two or even three alternative or simultaneous focused functions [...] The idea of a focused function, then, is vague.“ (Millikan 1984, 37)⁸⁵

⁸³ Ob diese Tendenz nach Millikan zu Recht besteht, ist weniger klar. Zumindest schreibt sie einigen Organen wie z.B. dem Herzen, nach diesem Muster eine Eigenfunktion zu, was auf eine positive Antwort hindeutet.

⁸⁴ Millikan zufolge finden sich biologische Beispiele für fokussierte Eigenfunktionen ausschließlich bezüglich des Nervensystems (Vgl. Millikan 1984, 36).

⁸⁵ Man kann sich aber dennoch des Eindrucks nicht ganz erwehren, dass von mehreren fokussierten Eigenfunktionen zu sprechen doch eine *contradictio in adjectio* ist.

Wir müssen die Vagheit bei der Bestimmung der Eigenfunktion einer Vorrichtung im Hinterkopf behalten und schauen, ob sich daraus nicht eventuell Schwierigkeiten ergeben für Millikans Lösungen des Problems der Missrepräsentation.

Eine zweite Schwierigkeit besteht m.E. in Millikans Redeweise davon, dass die HOREF-Mitglieder *Produkte* von FOREF-Mitgliedern sind bzw. durch diese *produziert* werden. „[...] similar genes produced similar products.“ (Millikan 1984, 22) Insbesondere in entwicklungsbiologischen Kontexten, wie der Ontogenese von Organen und anderen Strukturen, wird somit der Eindruck erweckt, als ob die DNA bzw. das Genom die Hauptursache dieser Entwicklungsprozesse ist. „Ex DNA omnia“ (Wolpert 1991, 77). Diese Auffassung privilegiert für Mahner und Bunge illegitimerweise einen einzigen Faktor aus einem ganzen Netzwerk von interagierenden Faktoren (Vgl. Mahner und Bunge 2000, 281). Für Mahner und Bunge gilt hingegen:

„Alle Entwicklungsprozesse von Biosystemen werden durch die systematische und gesetzmäßige Interaktion (der Elemente) seines Genoms, seiner extragenomischen Zusammensetzung und seiner Umgebung reguliert oder kontrolliert.“ (Mahner und Bunge 2000, 282)

Und daraus folgt:

„Es gibt kein exklusives (oder hinreichendes oder bevorzugtes) Kontrollsystem bei der Entwicklung der Organismen, wie z.B. sein Gensystem.“ (Mahner und Bunge 2000, 282)

In diesem Punkt ist Mahner und Bunge zuzustimmen, das Gensystem ist nur ein Faktor u.a. bei der Ontogenese und es ist definitiv nicht der alleinige Produzent des Phänotyps.

Die von Millikan vertretene direkte kausale Abhängigkeit bestimmter phänotypischer Merkmale, wie Organe oder Verhaltensweisen, von bestimmten Genen entspricht der „Gen für“-Redeweise vieler Genetiker. So kann für den Genetiker die Ursachenkette, die vom Genotyp zum Phänotyp führt, viele Glieder haben: der Biochemiker interessiert sich meist nur für die ersten Glieder, RNA und Protein; der Physiologe oder Anatom für spätere Glieder, wie Organe oder Gewebestrukturen; der Genetiker für Einzelorganismen für phänotypische Ausprägungen, wie Haarkrause oder Augenfarbe und schließlich interessiert sich der Verhaltensgenetiker für noch weiter entfernte Glieder, nämlich Verhaltensmuster, wie der Tanz der Honigbiene oder der Schnauzentriller des Stichelings (Vgl. Dawkins 2000, 245). In allen diesen Bereichen ist es üblich von Genen „für“ ein Verhalten, ein anatomisches Merkmal oder ein Protein zu reden und diese Redeweise hat durchaus ihre Berechtigung wie eine Argumentation von Dawkins deutlich macht (Vgl. Dawkins 2000, 22-31). Sobald wir ein Verhalten, wie z.B. das Löchergraben des Ameisenlöwen⁸⁶, (oder ein anatomisches Merkmal) als eine darwinistische Anpassung (siehe S. 64) betrachten, dann ist damit implizit auch Folgendes gesagt:

- a) in der Vergangenheit gab es eine Variation bezüglich des Verhaltens in der Population der Ameisenlöwen, d.h. einige Ameisenlöwen haben Löcher gegraben andere haben dieses Verhalten nicht gezeigt,
- b) diese Unterschiede im Verhalten waren vererbbar, d.h. nicht erlernt.

Daraus folgt nun, dass es zumindest in der Vergangenheit auch eine genetische Variation in der Ameisenlöwenpopulation gegeben haben muss, die der Variation im Verhalten entsprach.

⁸⁶ „Ameisenlöwen sind Insekten, Netzflüglerlarven, deren Aussehen und Gebaren an außerirdische Monster denken lassen. Es sind Räuber, die Gruben im lockeren Sand anlegen, und so lange geduldig verharren, bis Ameisen oder andere kleine Insekten in ihre fast perfekt kugelförmig angelegte Falle geraten. Deren Wände sind so steil, dass die Beutetiere nicht herausklettern können, wenn sie einmal in die Grube gestürzt sind. Der Ameisenlöwe sitzt unter dem Sand am Grund der Grube, wo er sich mit beängstigenden Zangen auf alles stürzt, was in das Loch fällt.“ (Dawkins, EP, 22)

D.h. die Mitglieder der Population müssen sich hinsichtlich bestimmter Allele unterscheiden haben. Allele sind dabei alternative Sequenzen eines beliebig langen Teilstücks auf den Chromosomen.

„Wenn wir auf einen Teil eines Chromosoms blicken, das fünf Cistrons lang ist, sind dessen Allele die alternativen Sätze von fünf Cistrons, die an übereinstimmenden Loci aller Chromosomen der Population vorhanden sind.“ (Dawkins 2000, 93)

Dem phänotypischen Unterschied entsprach ein genotypischer Unterschied:

Allel A	→	Löcher graben
Allel B	→	keine Löcher graben

In diesem Fall wäre das Allel A ein Gen „für“ Löchergraben. Diese Redeweise ist insofern legitim als sich die Genetiker in Bezug auf Anpassungen für bestimmte Unterschiede interessieren und „die Entwicklung des Embryos als Blackbox behandeln.“ (Dawkins 2000, 24)

„Weil Genetiker sich immer mit phänotypischen Unterschieden befassen, müssen wir uns nicht scheuen, Gene mit unendlich komplexen phänotypischen Auswirkungen vorauszusetzen – Auswirkungen, die sich nur unter hochkomplexen Entwicklungsbedingungen zeigen.“ (Dawkins 2000, 24)

Allerdings macht Dawkins an einem Beispiel auch sehr deutlich, dass die Gen-für-Redeweise oft nur einen höchst indirekten kausalen Zusammenhang ausdrückt (Vgl. Dawkins 2000, 27). „Gen für X“ heißt also weder, dass das Gen ein kausaler Hauptfaktor für die Hervorbringung von X sein muss, noch, dass das Gen X in irgendeiner Weise produziert. Von den zwei Honigbienenrassen, Brown und Van Scoy, zeigen erstere ein bestimmtes Hygiene-Verhalten, das Abdeckeln von Zellen, die kranke Brut enthält, und deren Entfernung nach dem Öffnen, und letztere nicht. Der Unterschied zwischen den zwei Rassen beruht nach dem Rothenbuhler-Modell⁸⁷ auf einem genetischen Unterschied an zwei Loci. Da Rothenbuhler aber gelegentlich beobachtete, dass auch Van-Scoy-Bienen das hygienische Verhalten zeigten, ist zu vermuten, dass beide die Brown- und die Van-Scoy-Bienen Vorfahren hatten, die das hygienische Verhalten ausübten.

„Und wahrscheinlich besitzen beide in ihrem Nervensystem den Mechanismus des Verhaltens für das Abdeckeln und Entfernen, nur haben die Van-Scoy-Bienen eben Gene, die verhindern, dass der Mechanismus aktiviert wird.“ (Dawkins 2000, 27)

Diese Gene sind nun in der Gen-für-Redeweise Gene für Nicht-Hygiene-Verhalten und deren Allele bei den Brown-Bienen sind dann die Gene für das Hygiene-Verhalten.

„Wenn wir noch weiter in der Zeit zurückgingen, würden wir vermutlich einen Vorgänger aller modernen Bienen finden, dem nicht nur das Hygieneverhalten fehlte, sondern der auch keine hygienischen Vorgänger gehabt hat. Es muss einen evolutionären Fortschritt gegeben haben, der dieses Abdeckungs- und Entfernenverhalten aus nichts aufgebaut hat, und bei diesem evolutionären Fortschreiten waren viele Gene einbezogen, die bei den Brown- als auch bei den Van-Scoy-Bienen fixiert sind.“ (Dawkins 2000, 27)

Der Ausdruck „Gen für Hygieneverhalten“ ändert also seine Referenz, je nach dem ob wir die längst ausgestorbenen Vorfahren der Brown- und Van-Scoy-Bienen (Fall 1) betrachten oder diese selbst (Fall 2). Dies können wir uns an einem Bild veranschaulichen:

⁸⁷ Benannt nach Rothenbuhler, der das hygienische Verhalten der Bienen untersucht hatte (1964).

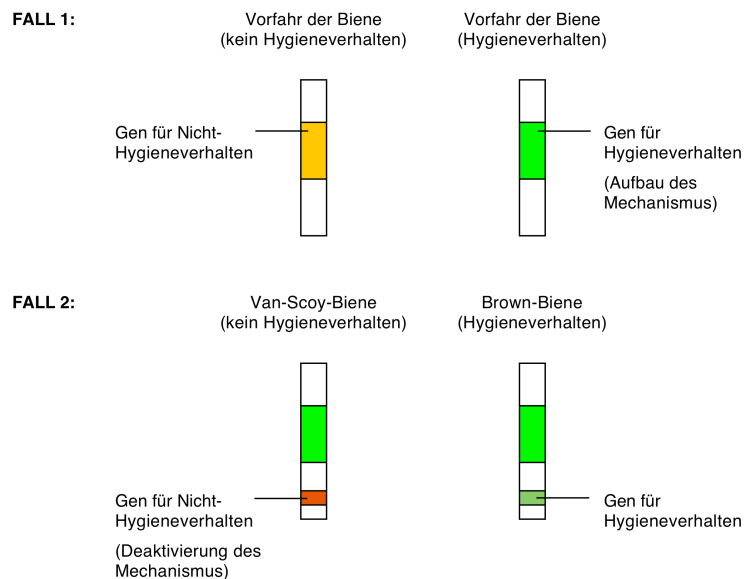


Abbildung 26: Kontextabhängigkeit der Referenz von „Gen für Hygieneverhalten“

Der zweite Fall zeigt deutlich, dass die Annahme der Ausdruck „Gen-für-X“ impliziere, dass das entsprechende Gen ein kausaler Hauptfaktor für die Hervorbringung von X ist, irreführend sein kann. Vor einem solchen Fehlschluss hat schon Gregory (1961) gewarnt:

„[...] die Entfernung irgendeines der weit verstreut angeordneten Widerstände könnte ein Radiogerät zum Pfeifen veranlassen. Aber daraus folgt nicht, dass die Pfeiftöne unmittelbar mit diesem Widerstand zusammenhängen, man kann daraus nur schließen, dass ein kausaler Zusammenhang lediglich höchst indirekt wirkt.“ (zitiert nach Dawkins 2000, 27)

Bei Millikans Ausführungen zur HOREF sollten wir also eher an Fall 1 als an Fall 2 denken. Aber selbst dann stehen uns momentan in den allermeisten Fällen keine Normalen Erklärungen zur Verfügung, wie ein bestimmter Teil des Genoms (oder besser wie unter dessen Beteiligung) die ontogenetische Entwicklung einer bestimmten phänotypischen Ausprägung bewirkt. Mag es in Zukunft solche Erklärungen geben und die Fortschritte von im Bereich der Evolutionären Entwicklungsbiologie (Evolutionary Developmental Biology kurz: Evo-Devo) sprechen dafür, jetzt gibt es sie jedenfalls nicht. Folglich gilt die Zugehörigkeit zu einer HOREF oft nur unter Vorbehalt, dies betrifft auch Millikans Lieblingsbeispiele wie Bientänze, das Paarungsverhalten von Stichlingen und Herzen.

Kehren wir nun zur Erläuterung des Schemas (Abbildung 21) zurück. Wir nehmen jetzt einfach an die Proteine sind irgendwie für die Bildung der Organe mitverantwortlich, vor allem soll aus der Ähnlichkeit der Proteine eine Ähnlichkeit der Organe folgen. Die Organe a, b und c wären dann also die Produkte von Mitgliedern einer REF zweiter Ordnung deren Funktion es ist solche Produkte hervorzubringen und die Produktion sollte einer Normalen Erklärung zugänglich sein. Somit bilden die Organe a, b und c eine REF dritter Ordnung. Die Organe a, b und c hätten dann wieder nahezu gleiche Eigenschaften und unter diesen Eigenschaften mag eine Eigenschaft C3 sein, die einen fitnesserhöhenden Effekt von C3-Trägern gegenüber Nicht-C3-Trägern hat. C3 soll hier eine Eigenschaft sein, die unter bestimmten Bedingungen gelegentlich zur Ausübung eines bestimmten Verhaltens führt. Das gelegentliche Hervorbringen des Verhaltens wäre dann die Eigenfunktion (F3) der Organe a, b und c. Die Verhaltensausübungen a, b und c wären dann die Produkte von Mitgliedern einer REF dritter Ordnung und sie selbst bilden eine REF vierter Ordnung. Wieder beruht die Ähnlichkeit der Verhaltensausübungen auf der Ähnlichkeit der Organe und wieder mag eine Eigenschaft C4 unter den Eigenschaften des Verhaltens sein, die ihren Trägern einen

Selektionsvorteil verschafft, weil sie einen bestimmten Effekt bewirkt. Diesen Effekt zu bewirken ist dann die Eigenfunktion des Verhaltens (F4).

Der Teil des Schemas, der die Hervorbringung des Verhaltens durch das Organ betrifft, müsste eigentlich anders dargestellt werden. In etwa so:

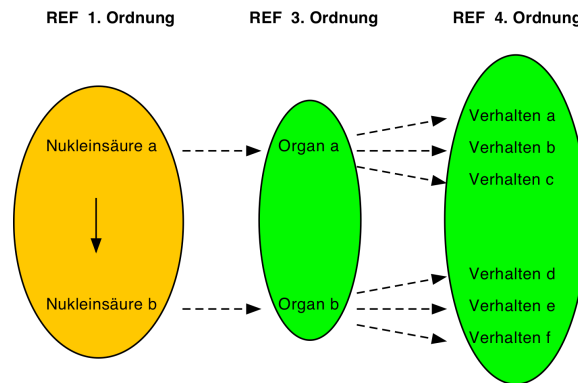


Abbildung 27: Schematische Darstellung des Zusammenhangs zwischen FOREFs und HOREFs (Verhalten)

Denn ein und dasselbe Organ bzw. ein und dieselbe physiologische Struktur kann dasselbe Verhalten oft mehr als einmal hervorbringen. Der Mechanismus der dem Schnauzentriller des Stiehls zugrunde liegt, bringt für die Dauer seiner Existenz mehr als einen Schnauzentriller hervor.

Vom letzten Schema ausgehend können wir uns nun auch den HOREFs zuwenden, die durch die zweite Definition bestimmt sind. Das Schema muss dazu nur leicht modifiziert werden, denn nur diejenigen Verhaltensäußerungen sollen zur HOREF gehören, die den vorangegangenen Verhaltensäußerungen gleichen. Auf diese Weise will Millikan auch individuell erlernte Verhaltensweisen erfassen, wie diejenigen, die durch klassische oder instrumentelle Konditionierung entstehen. Dies ergibt sich eher aus ihren Erläuterungen als aus der Definition selbst.

„Learned behavior resulting from training or a trail-and-error procedure are members of higher-order reproductively established families in accordance with (2) [gemeint ist die zweite Definition von HOREFs], assuming that it is a proper function of the mechanisms that produce such behaviors to turn out again the *same* behavior as was rewarded or found to be successful earlier.“ (LTOBC, 25)

D.h., das Organ (wahrscheinlich eher ein ganzer Komplex von Organen inklusive Teile des Nervensystems) bringt zunächst mehr oder weniger zufällig eine Reihe verschiedener Verhaltensausübungen hervor (trail and error), bis schließlich eine diese Verhaltensausübungen „belohnt“ wird (reinforcement) und fortan beibehalten wird. Die gelernten Verhaltensäußerungen ähneln sich insofern, als das Organ, das sie hervorbringt auch einen Lernmechanismus enthält, der eben genau dafür sorgt, dass die gelernte Verhaltensreaktion wiederholt wird.

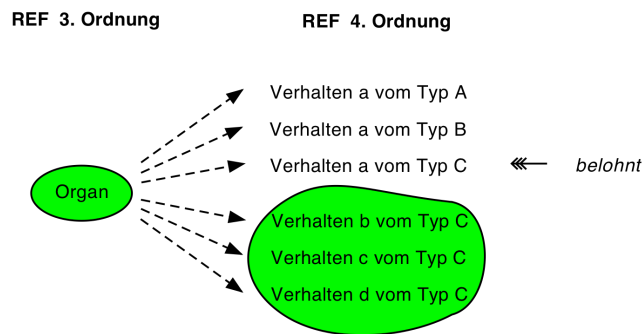


Abbildung 28: Schematische Darstellung des Zusammenhangs zwischen zwei HOREFs (konditioniertes Verhalten)

Wie bei der Ontogenese der Organe stellt sich hier wieder die Frage, ob uns bezüglich der klassischen oder der instrumentellen Konditionierung bei irgendeinem Organismus schon so etwas wie Millikans Normale Erklärung zur Verfügung steht. Für einfachere Formen des Lernens, wie Habituation und Sensitivierung, gibt es bereits Normale Erklärungen (Vgl. Kandel 1976)⁸⁸. Aber bei diesen Lernformen wird gerade nicht dasselbe Verhalten bezüglich desselben Stimulus wiederholt. Bei Habituation nimmt die Intensität des Verhaltens bei gleicher Reizstärke ab, während bei der Sensitivierung die gleiche Verhaltensreaktion auf einen viel schwächeren Reiz folgt.

Insgesamt bleibt also bei Vielem (Organe, Verhalten), was Millikan als HOREF betrachtet, aufgrund des Fehlens einer Normalen Erklärung für deren Produktion zunächst offen, ob es sich dabei tatsächlich um HOREFs handelt.

⁸⁸ Dies gilt insbesondere für Verhaltensweisen des Kalifornischen Seehasen (*Aplysia californica*).

2.1.1.3.6 Relationale Eigenfunktionen

In diesem Abschnitt soll Millikans Definition einer relationalen Eigenfunktion eingeführt und erläutert werden. Im Zusammenhang mit den relationalen Eigenfunktionen erweitert Millikan das Vokabular ihrer Theorie um folgende weitere Begriffe:

- adaptierte⁸⁹ Eigenfunktion (adapted proper function)
- adaptierte Vorrichtung (adapted device)
- abgeleitete Eigenfunktion (derived proper function)
- Adaptor (adaptor)

Diese Erweiterung soll es ermöglichen auch flexibles Verhalten und damit einzelne gänzlich neue Verhaltensäußerungen, d.h. bei noch keinem Vorfahren aufgetretene Verhaltensäußerungen, im Rahmen ihrer Theorie zu erfassen.⁹⁰

Ausgangspunkt von Millikans Überlegungen ist der Begriff der relationalen Eigenfunktion. Das Haben einer relationalen Eigenfunktion bestimmt sie wie folgt:

„A device has a relational proper function if it is its function to do or to produce something that bears a specific relation to something else.“ (Millikan 1984, 39)

Was damit gemeint sein soll, erläutert Millikan an drei Beispielen: dem Tarnverhalten des Chamäleons, der Chemotaxis von Amöben und dem Tanzverhalten der Honigbiene.

Das Chamäleon besitzt eine Vorrichtung (pigment arranger), deren Eigenfunktion es ist, die Hautfarbe des Chamäleons an die Farbe des Untergrunds, auf dem es sich befindet, anzupassen, um das Chamäleon für Fressfeinde weniger sichtbar zu machen. Die Funktion des *pigment arrangers* ist es also etwas zu produzieren, nämlich die Hautfarbe, das in einer spezifischen Relation, nämlich „gleiche Farbe wie“, zu etwas anderem, nämlich dem Untergrund, steht.

Amöben besitzen eine Vorrichtung⁹¹, deren Funktion es ist ein Schwimmverhalten der Amöben in Abhängigkeit von chemischen Gradienten (z.B. Zuckergehalt) zu produzieren, um

⁸⁹ Auch diese Wortwahl ist aufgrund der Assoziationen zu „Adaptation“ in der Evolutionstheorie etwas unglücklich, wie Millikan in einer Fussnote selbst zugesteht: „This term is not ideal. Especially, it calls to mind evolutionary adaptation which is a long process rather than an affair of the moment. The primary meaning of ‚adapt‘ is simply to make fit or suitable, to adjust. Unfortunately none of these synonyms have graceful forms parallel to all of those that will be needed of ‚adapt.‘ Besides, there is a special similarity between the meaning of ‚adapted‘ as I will use it in the context ‚adapted device‘ and the evolutionary sense of ‚adapted.‘ In each case we have something that is both determined by and fitted to its environment such as to serve, in that environment, some invariant function-ultimately, in each case, survival or reproduction.“ (Millikan 1984, 336).

⁹⁰ Dies ist insbesondere in Bezug auf Millikans Fernziel relevant eine biologische Fundierung von menschlichem Sprechen und Denken zu liefern. Hintergrund der Erweiterung dürfte ein generelles Bedenken sein, das mit dem stark evolutionstheoretisch geprägten Funktionsbegriff von Millikan zusammenhängt. Das Bedenken geht etwa in folgende Richtung: Es ist nahezu unmöglich den Selektionsvorteil einzelner Überzeugungen plausibel zu machen, zudem haben einige unserer Gedanken ganz und gar neue Gehalte und können folglich nicht durch eine Selektionsgeschichte erklärt werden, also sind Millikans Ansatz von vornherein sehr enge Grenzen gesteckt. Millikan versucht in diesem Kapitel die begrifflichen Grundlagen zu legen, um einem solchen Einwand zu entgehen. Der Einwand wird aber trotz des zweiten Kapitels gegen die Biosemantik vorgetragen (Vgl. z.B. Schlicht 2008, 66), wahrscheinlich in Unkenntnis dessen Inhalts.

⁹¹ Millikan behauptet hier eigentlich, dass die gesamte Amöbe die Vorrichtung sei. Sie schreibt also offenbar auch ganzen Organismen Eigenfunktionen zu, was Wild in seiner Darstellung nicht ganz korrekt wiedergibt (Vgl. Wild 2010, 51f). „Organismen haben [...] keine Echten Funktionen [proper functions, A.W.]“ (Wild 2010, 52). In der Sache hat Wild hier aber m.E. Recht, ganzen Organismen Eigenfunktionen zuzuschreiben ist merkwürdig und diesen Beispiel vermutlich einfach falsch. Die

es aus schädlichen Umgebungen und hin zu günstigen zu manövrieren. Die Funktion dieses chemotaktischen Mechanismus' ist es also etwas zu produzieren, nämlich ein Schwimmverhalten, das in einer spezifischen Relation, nämlich „in Richtung von“, zu etwas anderem, nämlich der Zuckerquelle, steht.

Bienen besitzen eine Vorrichtung (*dance choreographing device*), deren Eigenfunktion es ist, den Tanz einer von einer Nektarquelle heimkehrenden Biene auf dem Ort der Quelle abzustimmen, damit andere Bienen ebenfalls zu dieser Nektarquelle finden und letztlich möglichst viel Nektar zum Bienenstock gebracht wird. Die Funktion des *dance choreographing device* ist es also etwas zu produzieren, nämlich ein Tanzverhalten, dessen Elemente in spezifischen Relationen⁹² zu etwas anderem, nämlich der Nektarquellen, stehen.

Um das Besondere von relationalen Eigenfunktionen zu erfassen, ziehen wir zu den drei genannten Beispielen noch ein weiteres Beispiel den Kiemen-Einzieh-Reflex von *Aplysia* hinzu.

Das Nervensystem der Meeresschnecke *Aplysia* ist eines der am besten untersuchten Nervensysteme überhaupt und der Kiemen-Einzieh-Reflex ist das einfachste Verhalten, das *Aplysia* zeigt (Vgl. Kandel 2006, 188). Eric Kandel beschreibt das Verhalten folgendermaßen:

„The gill is an external organ that *Aplysia* uses to breathe. It lies in a cavity of the body wall called the mantle cavity and is covered by a sheet of skin called the mantle shelf. The mantle shelf ends in the siphon, a fleshy spout that expels seawater and waste from the mantle cavity. Touching the siphon lightly produces a brisk defensive withdrawal of both siphon and the gill into the mantle cavity. The purpose of the withdrawal reflex is clearly to protect the gill, a vital and delicate organ, from possible damage.“ (Kandel 2006, 188f).

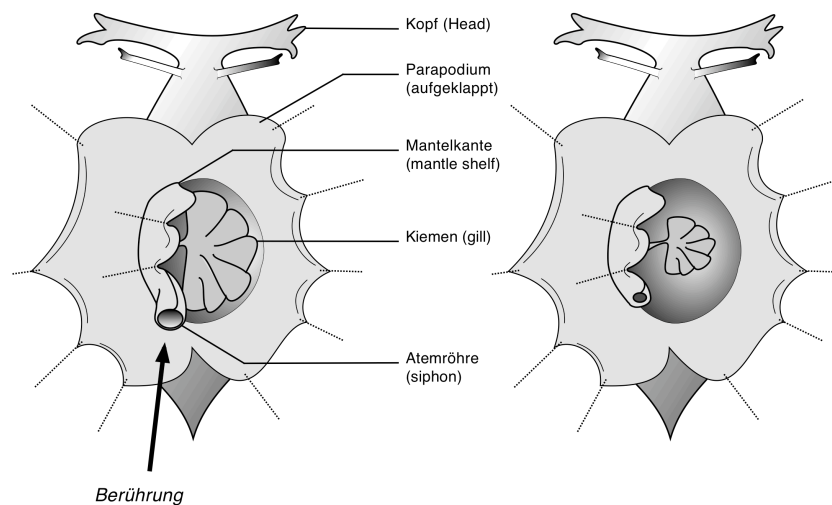


Abbildung 29: Der Kiemen-Einzieh-Reflex bei *Aplysia* (In Anlehnung an Kandel 1976, 351)

Chemotaxis ist insbesondere beim Bakterium *Escherichia coli* sehr gut untersucht und man hat hier auch die Strukturen und Komponenten identifiziert, die der Chemotaxis zugrunde liegen (Vgl. Baker et al 2005). Diese Strukturen entsprechen keinesfalls dem ganzen Bakterium und es ist zu vermuten, dass dasselbe für Amöben gilt. Ich weiche hier dementsprechend von Millikans Darstellung ab.

⁹² Die genauen Relationen sind für den sogenannten Schwänzeltanz recht kompliziert und sollen hier nicht dargestellt werden (siehe aber Abschnitt 2.1.1.3.8). Der Bientanz kann wohl als Millikans Lieblingsbeispiel gelten und ist zwar auch eine relationale Eigenfunktion, aber zudem noch mehr, insofern ist Millikans verfrühtes Anführen dieses Beispiels etwas unglücklich.

In jedem der vier Beispiele liegt eine ähnliche Struktur vor: Eine HOREF hat unter bestimmten Normalen Bedingungen die Eigenfunktion eine weitere HOREF zu produzieren, die wiederum die bestimmte Eigenfunktion hat einen bestimmten Effekt bzw. eine weitere HOREF zu produzieren (siehe oberste Zeile in Abbildung 30). Diesen Zusammenhang können wir für die vier Beispiele wie folgt schematisch darstellen:

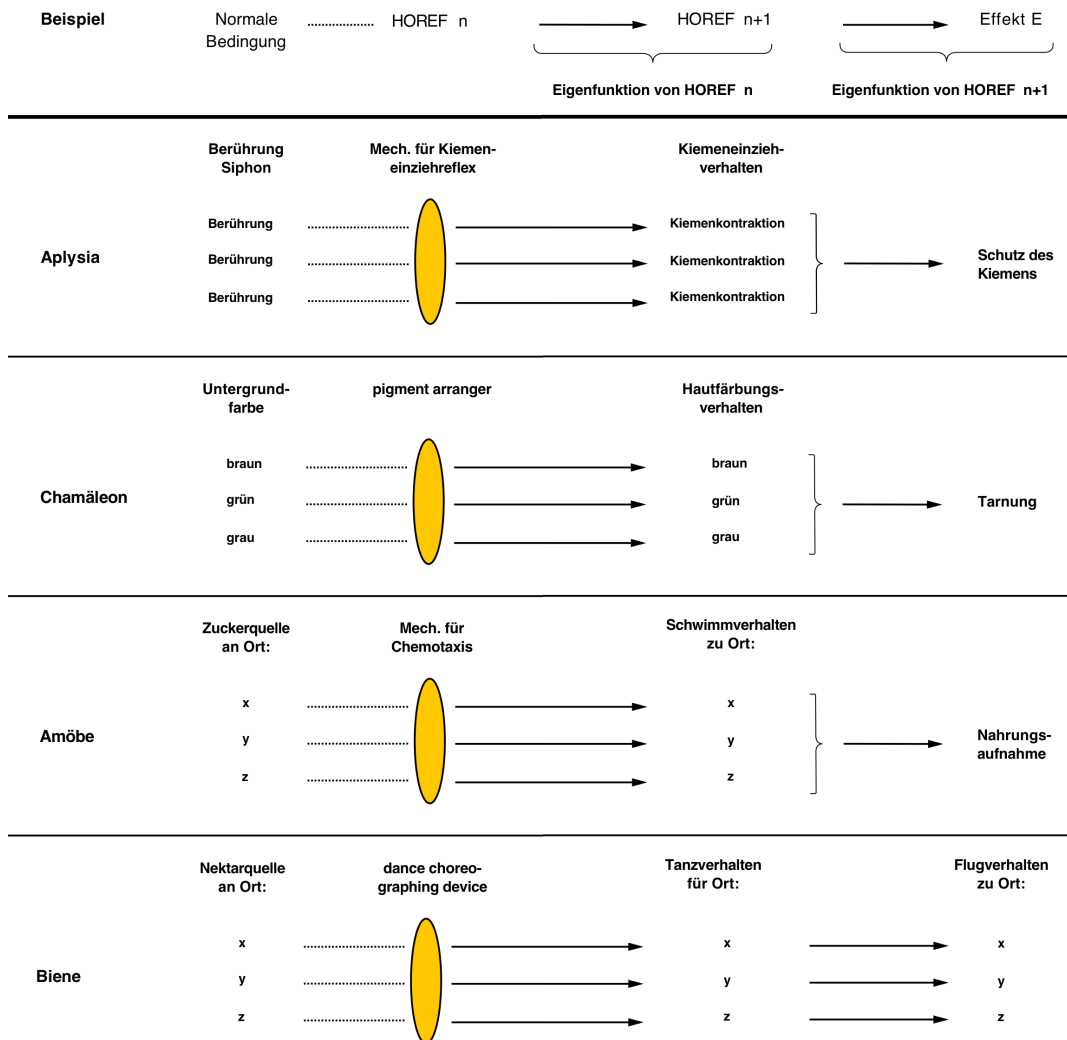


Abbildung 30: Die Struktur relationaler Eigenfunktionen anhand von Beispielen

In den Beispielen von Chamäleon, Amöbe und Biene gibt es Variationen unter den Normalen Bedingungen, denen im Sinne einer mathematischen Abbildungsfunktion Variationen im Verhalten zugeordnet sind. Im Schema sind immer nur jeweils drei Variationspaare angegeben, diese stehen jedoch für eine nahezu unendliche Anzahl an Variationspaaren. Im Aplysia-Beispiel besteht ebenfalls eine Zuordnung zwischen Normaler Bedingung und Verhalten aber diese weist im Gegensatz zu den anderen Beispielen keine Variationen auf.⁹³ Man könnte versucht sein neben der Zuordnung (Berührung – Kiemenkontraktion) auch die Zuordnung (keine Berührung – keine Kiemenkontraktion) als Variation des Kiemen-einziehreflex-Mechanismus in Betracht zu ziehen, dann gebe es zwar jeweils nur zwei Varianten, aber es bestünde kein prinzipieller, sondern nur ein quantitativer Unterschied zu

⁹³ Wir ignorieren hier Variationen in Berührungsintensität und Kiemenkontraktion, ebenso wie das Lernverhalten von Aplysia.

den anderen Beispielen. So gesehen, lassen sich kaum Verhaltenweisen finden, deren Produzieren keine relationalen Eigenfunktionen wären. Eventuell wäre der *random walk* eines Bakteriums eine solche⁹⁴. Insgesamt legen Millikans Beispiele jedoch nahe, dass nur diejenigen Vorrichtungen relationale Eigenfunktionen haben, bei denen mehr als zwei Umweltvarianten mehr als zwei Verhaltensvarianten zugeordnet sind⁹⁵.

Die Zusammenhänge bei relationalen Eigenfunktionen lassen sich sowohl allgemein als auch konkret beschreiben. Für die konkrete Beschreibung benötigt Millikan die bereits erwähnten Begriffe: adaptierte Eigenfunktion (adapted proper function), adaptierte Vorrichtung (adapted device) und abgeleitete Eigenfunktion (derived proper function). Wir können nun zunächst die allgemeine Beschreibung der konkreten gegenüberstellen und uns anschließend anhand unserer Beispiele klar machen, worauf die jeweiligen Begriffe referieren.

Allgemein	Konkret
Relationale Eigenfunktion (direkte Eigenfunktion)	Adaptierte Eigenfunktion (keine direkte Eigenfunktion)
Produkt der relationalen Eigenfunktion (HOREF)	Adaptierte Vorrichtung (keine HOREF)
Funktion des Produkts der relationalen Eigenfunktion (direkte Eigenfunktion)	Abgeleitete Eigenfunktion (keine direkte Eigenfunktion)

Abbildung 31: Zuordnung von Begriffen der allgemeinen und der konkreten Beschreibung von relationalen Eigenfunktionen

Adaptierte Eigenfunktion: Eine Vorrichtung, die eine relationale Eigenfunktion hat, erwirbt unter ganz konkreten Bedingungen eine adaptierte Eigenfunktion.

„When a device has a relational proper function, *given* some specific thing that the device is now supposed to produce in relation to, the device acquires what I will call a *adapted proper function*.“ (Millikan 1984, 40)

	Relationale Eigenfunktion (direkte Eigenfunktion)	Adaptierte Eigenfunktion (keine direkte Eigenfunktion)
Chamäleon	Produzieren einer Hautfärbung, die Untergrund gleicht	Produzieren von brauner Haut- färbung bei tatsächlich vorhandenen braunen Untergrund
Amöbe	Produzieren von Schwimmverhalten zum Ort der Zuckerquelle	Produzieren von Schwimmverhalten an Ort x bei tatsächlich vorhandener Zuckerquelle an Ort x
Biene	Produzieren von Tanzverhalten passend zu Ort der Nektarquelle	Produzieren von Tanzverhalten passend zu Ort x bei tatsächlich vorhandener Nektarquelle an Ort x

Abbildung 32: Relationale und adaptierte Eigenfunktion bei unseren Beispielen

Adaptierte Eigenfunktionen sind keine direkten Eigenfunktionen, weil die Vorfahren der Funktionsträger nicht aufgrund des spezifischen Effekts der adaptierten Eigenfunktion

⁹⁴ Das Bakterium bewegt sich dabei ohne dass die Richtung der Bewegung durch irgendeine Umweltbedingung bestimmt wird. Den Ort zu wechseln ist statistisch gesehen besser als sich nicht gar zu bewegen, hierin liegt der Vorteil dieses Verhaltens.

⁹⁵ Eine Zwei-zu-zwei-Zuordnung wie sie eventuell das Aplysia-Beispiel darstellt, könnten wir dann als primitiven Grenzfall einer relationalen Eigenfunktion ansehen.

überlebt haben. Es war nicht diese ganz spezielle braune Hautfärbung, nicht das Schwimmverhalten an diesen bestimmten Ort und nicht dieser spezifische Tanz, die dafür gesorgt haben, dass die Vorfahren der Funktionsträger einen Selektionsvorteil hatten.

Adaptierte Vorrichtung: Das Produkt der adaptierten Eigenfunktion bezeichnet Millikan als adaptierte Vorrichtung (adapted device). Nach Millikan bilden adaptierte Vorrichtungen keine HOREFs⁹⁶ im Extremfall können sie sogar noch nie in der Vergangenheit vorgekommen sein⁹⁷.

	Produkt der relationalen Eigenfunktion (HOREF)	Adaptierte Vorrichtung (keine HOREF)
Chamäleon	Hautfärbung, die Untergrund gleich	braune Hautfärbung bei tat- sächlich vorhandenen braunen Untergrund
Amöbe	Schwimmverhalten zum Ort der Zuckerquelle	Schwimmverhalten an Ort x bei tatsächlich vorhandener Zuckerquelle an Ort x
Biene	Tanzverhalten passend zu Ort der Nektarquelle	Tanzverhalten passend zu Ort x bei tatsächlich vorhandener Nektarquelle an Ort x

Abbildung 33: Produkt der relationalen Eigenfunktion und adaptierte Vorrichtung bei unseren Beispielen⁹⁸

Dasjenige, woran die adaptierte Vorrichtung adaptiert ist, bezeichnet Millikan als **Adaptor** (Vgl. Millikan 1984, 40). Es kann vorkommen, dass eine adaptierte Vorrichtung zugleich an eine andere adaptierte Vorrichtung adaptiert ist und damit an den Adaptor dieser zweiten adaptierten Vorrichtung. Das ist beim bereits erwähnten Flugverhalten der Honigbiene der Fall. Das Flugverhalten ist nach Millikan eine adaptierte Vorrichtung, die zunächst an den Bientanz der heimkehrenden Biene adaptiert ist. Der Tanz ist aber seinerseits eine adaptierte Vorrichtung, deren Adaptor eine bestimmte Nektarquelle ist. Den ultimativen Adaptor, hier die Nektarquelle, bezeichnet Millikan als *original adaptor*.

„The furthest removed thing to which a device is adapted is that to which it is originally adapted – its original adaptor.“ (Millikan 1984, 41)

Abgeleitete Eigenfunktion: Die Eigenfunktion einer adaptierten Vorrichtung leitet sich aus der Eigenfunktion der Vorrichtung ab, die sie produziert hat, und zwar aus der Eigenfunktion, die über die unmittelbare Produktion der adaptierten Vorrichtung hinausgeht.

„The proper functions of adapted devices are derived from proper functions of the device that produces them that lie *beyond* the production of these adapted devices themselves. I will call the proper functions of adapted devices *derived proper functions*.“ (Millikan 1984, 41)

⁹⁶ Der Grund scheint zu sein, dass sie nicht die Produkte von direkten Eigenfunktionen sind, was allerdings nur unter bestimmten Beschreibungen der Fall ist. Abgesehen von dem genannten Extremfall könnten adaptierte Vorrichtungen m.E. durchaus HOREFs bilden.

⁹⁷ D.h. es könnte ein ganz spezifisches Hautfärbungsmuster oder ein Bientanz produziert werden, der noch nie in der Geschichte der Spezies produziert wurde. Dies ist zwar unwahrscheinlich aber durchaus möglich.

⁹⁸ Es ist terminologisch sicher etwas unglücklich das Schwimm- und Tanzverhalten als Vorrichtungen zu bezeichnen. Es ist aber auch kein Problem, da *adapted device* ein terminus technicus ist.

Was ist mit „Eigenfunktionen, die über die Produktion der adaptierten Vorrichtung hinausgeht“ gemeint. Wir hatten bereits festgestellt, dass Funktionen zunächst selektierte Effekte sind. Effekte können weitere Effekte auslösen, so dass die Ursache des ersten Effekts indirekt auch die Ursache des zweiten Effekts ist. Ebenso kann der Träger einer Eigenfunktion durch deren Ausübung einen weiteren Effekt auslösen, der auch als dessen Eigenfunktion gelten kann (siehe S. 75-77). Die Eigenfunktion einer für ein bestimmtes Protein codierenden RNA ist die Synthese des Proteins, aber darüber hinaus auch dasjenige, was das Protein bewirkt, z.B. die Katalyse der RNA selbst.

	Funktion des Produkts der relationale Eigenfunktion (direkte Eigenfunktion)	Abgeleitete Eigenfunktion (keine direkte Eigenfunktion)
Chamäleon	Tarnung (invariant)	Tarnung (invariant)
Amöbe	Nahrungsaufnahme (invariant)	Nahrungsaufnahme (invariant)
Biene	Nektar in Stock (invariant)	Nektar in Stock (invariant)
	Flugverhalten in Abhängigkeit von Tanz (adaptiert)	Flugverhalten in Abhängigkeit von Nektarquelle (adaptiert)

Abbildung 34: Funktion des Produkts der relationalen Eigenfunktion und abgeleitete Eigenfunktion bei unseren Beispielen

Hinsichtlich der abgeleiteten Eigenfunktionen unterscheidet Millikan noch einmal zwischen invarianten abgeleiteten Eigenfunktionen (invariant derived proper function) und adaptierten abgeleiteten Eigenfunktionen (adapted derived proper function). Eine adaptierte abgeleitete Eigenfunktion liegt vor, wenn auch der Effekt, der über die Produktion der adaptierten Vorrichtung hinausgehenden Eigenfunktion des Vorrichtungsproduzenten als an einen Adaptor⁹⁹ angepasst beschrieben werden kann. Das klingt zunächst verwirrend. Ein solcher Fall liegt auch nur beim Bienenbeispiel vor, das, wie bereits gesagt, deutlich komplexer ist als die beiden anderen Beispiele und eigentlich einen Vorgriff darstellt. Der Effekt des *dance choreographing devices*, der über die Produktion des Tanzes hinausgeht, wäre hier das Auslösen eines bestimmten Flughaltens bei anderen Bienen und zwar ein Flugverhalten zu einer bestimmten Nektarquelle, die auch schon der Adaptor für den Tanz gewesen ist. Wie wir in Abbildung 30 sehen können weicht in diesem Fall die Beschreibung der direkten Eigenfunktion des Produktes der relationalen Eigenfunktion (d.h. der direkten Eigenfunktion des Bientanzes) von der Beschreibung der abgeleiteten Eigenfunktion der adaptierten Vorrichtung (d.h. adaptierten abgeleiteten Eigenfunktion des Bientanzes) ab. Hieraus ergibt sich die Möglichkeit konfligierender Eigenfunktionen, die wiederum eine wichtige Rolle bei Millikans Lösungsansätzen zum Problem der Missrepräsentation spielen. Wir werden also später auf diese Möglichkeit zurückkommen (unter Abschnitt 2.2.3.2).

Abschließend sei noch eine kurze Bemerkung zu den Normalen Erklärungen von adaptierten und abgeleiteten Eigenfunktionen erlaubt. Da Normale Erklärungen angeben, *wie* die Vorfahren einer Vorrichtung in der Vergangenheit eine Funktion ausgeübt haben und sowohl adaptierte als auch abgeleitete Eigenfunktionen gänzlich neu sein können, ist unklar, worin deren Normale Erklärungen bestehen sollen. Millikan zufolge sind Normale Erklärungen von adaptierten und abgeleiteten Eigenfunktionen einfach die Normalen Erklärungen der entsprechenden direkten Eigenfunktionen. Daraus folgt, dass es sich auch bei diesen Erklärungen immer um generelle Erklärungen handelt.

⁹⁹ Der Adaptor ist dann derselbe wie derjenige der adaptierten Vorrichtung.

„An adapted proper function is a relational proper function adapted to a given context. If the relational proper function is performed Normally, clearly the adapted function is performed Normally too. The Normal explanation for proper performance of an adapted proper function is thus a general explanation that tells how it happens that the device produces or does things that bear certain relations to its adaptors.“ (Millikan 1984, 43f)

2.1.1.3.7 **Standardisierende und stabilisierende Eigenfunktionen**

In diesem Abschnitt soll Millikans Auffassung von standardisierenden und stabilisierenden Eigenfunktionen eingeführt und erläutert werden. Millikans Ausführungen hierzu sind weniger klar als zu den relationalen Eigenfunktionen. Dies liegt daran, dass sie keine allgemeine Definition von standardisierenden und stabilisierenden Eigenfunktionen gibt. Stattdessen erläutert Millikan das, was wir unter solchen Funktionen verstehen sollen, nahezu ausschliesslich anhand menschlicher Kommunikation¹⁰⁰. So bestimmt sie lediglich die standardisierende und stabilisierende Eigenfunktion von sprachlichen Ausdrücken (language device):

„The stabilizing and standardizing proper function (or functions) of a language device is that hypothesized function (or functions) that tends at the same time to keep speakers using the device in standard ways and to keep hearers responding to it in standard ways, thus stabilizing its function or functions.“ (Millikan 1984, 31f)

Der Bezug auf die menschliche Kommunikation ist für die naturalistischen Ansprüche von Millikans Theorie problematisch, denn die Erklärung des standardisierenden und stabilisierenden Effekts sprachlicher Ausdrücke scheint Absichten und Zwecke seitens der Sprecher und Hörer vorauszusetzen. So würden Sprecher sicherlich aufhören Aufforderungen zu äußern, wenn diese nie befolgt werden würden.¹⁰¹ Warum Sprecher so handeln würden, wird aber von Millikan nicht explizit erläutert. Vielleicht weil es zu trivial erscheint: Die Sprecher begreifen einfach nach einer Weile, dass das Äußern einer Aufforderung sinnlos ist, weil sie damit nicht erreichen, was sie bezwecken. Dies bedeutet aber Folgendes: Um zu verstehen, was ein Zeichen ist, muss man verstehen, was eine standardisierende und stabilisierende Eigenfunktion ist und um diese zu verstehen, müssen wir kognitive Fähigkeiten, wie das Verfolgen von Zwecken und Absichten voraussetzen. Letzteres untergräbt die naturalistischen Ansprüche der Theorie. Wir dürfen ja nicht vergessen, dass es uns darum geht, mit Hilfe von Millikans Theorie dem Homunkulus-Einwand zu entgehen. Insofern ist es fragwürdig, standardisierende Eigenfunktionen an einem Beispiel zu erläutern, in dem kognitive Fähigkeiten vorausgesetzt werden müssen. Millikan scheint diesen Einwand zu antizipieren, wenn sie darauf hinweist, dass die Zwecke der Sprecher und Hörer nicht immer in expliziten Absichten bestehen müssen.

„When a language device serves its stabilizing function in accordance with a Normal explanation, the speaker and hearer both purpose this function. But we may not have to think of this ‚purposing‘ as corresponding, always, to explicit intentions. It may be enough that the stabilizing function of the language device be the same as a derived proper function of the speaker's utterance and of the hearer's response mechanism or program.“ (Millikan 1984, 53)

¹⁰⁰ Kurze Erwähnung findet der Paarungstanz des Stichelings. Millikan geht aber nicht näher auf das Beispiel ein (Vgl. Millikan 1984, 3).

¹⁰¹ Millikan drückt dies etwas allgemeiner aus: „Hearers who are too often manipulated in ways that go against their own purposes will stop responding to language as they once would have done. Just as speakers will continue to speak only if there is some correlation between what they say and how hearers respond, so hearers will continue to listen and to follow certain patterns of translation into understanding and sometimes into belief or intention and action only if these responses are of some use in relation to the hearer's own projects in the world of use at least in some critical proportion of cases. And if hearers stop listening or stop responding in ways that show some uniformities, speakers in turn will stop speaking. Hence one of the proper functions of the various language devices must be to do things that keep hearers responding in the old ways with some degree of uniformity. Otherwise the language device would die out.“ (Millikan 1984, 31).

Da Millikan aber auf die angesprochenen Mechanismen nicht näher eingeht, bleiben diese reine Spekulation und können den Einwand kaum entkräften.

Es ist aber meines Erachtens durchaus möglich Millikans Auffassung einer standardisierenden Eigenfunktion auch ohne die Orientierung an menschlicher Kommunikation zu verstehen. Zunächst können wir uns die grundlegende Struktur klar machen, in der standardisierende Eigenfunktionen vorkommen können. Die Grundstruktur besteht aus zwei Vorrichtungen A und B, die in folgender Weise voneinander abhängig sind: Die Eigenfunktion von A besteht darin, den Effekt E zu produzieren. Der Effekt E wiederum ist eine Normale Bedingung dafür, dass B seine Eigenfunktion, den Effekt F hervorzubringen, ausüben kann. Der Effekt F wirkt sich dann positiv auf die Erhaltung und Vermehrung von A und B aus.

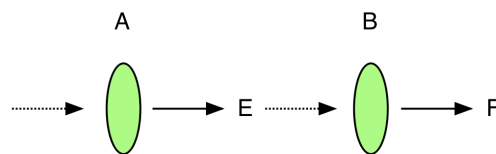


Abbildung 35: Grundstruktur standardisierender Eigenfunktionen

D.h., weder A noch B können ihre Eigenfunktionen ohne den jeweils anderen ausüben. Standardisierende Eigenfunktionen können nur da auftreten, wo Träger von Eigenfunktionen wie in Abbildung 35 „in Reihe geschaltet“ sind.

Der standardisierende Aspekt kommt zum Tragen, wenn dem verbindenden Element (Effekt E) eine gewisse Arbitrarität anhaftet.

Worin diese Arbitrarität bestehen soll, lässt sich mit Hilfe spieltheoretischer Überlegungen erläutern. Betrachten wir ein einfaches Koordinationsspiel¹⁰². Zwei Autofahrer befinden sich auf den gegenüberliegenden Enden einer Brücke, die gerade so breit ist, dass die Autos nebeneinander Platz finden. Beide Fahrer wollen die Brücke passieren, können aber nicht miteinander kommunizieren. Die Fahrer haben jeweils die Optionen, ganz links oder ganz rechts zu fahren. Das Spiel hat dementsprechend vier mögliche Ausgänge (siehe Abbildung 36), von denen wir nur die sicheren Überquerungen als wünschenswert betrachten.

¹⁰² Schon David Lewis hatte in *Konventionen* (1975) versucht die Arbitrarität von sprachlichen Zeichen mit Hilfe von Koordinationsspielen zu erklären. Millikan begründet in einer Fussnote kurz, warum sie nicht näher auf Lewis eingeht: “The reader familiar with David Lewis’s *Conventions* (Harvard University Press, Cambridge, Mass., 1969) may be troubled that I do not relate the ideas in these two paragraphs to Lewis’ work. In Chapter 3 I will argue that language is not conventional in the way Lewis describes, for Gricean intentions do not drive language use in the ordinary case. These arguments will be broad enough to constitute a critique of Lewis’s analysis of conventions generally.” (Millikan 1984, 335) Die Argumentation gegen Grice kann hier unberücksichtigt bleiben, denn sie ist nur für Millikans höhere Ansprüche (Vgl. Abschnitt 1.4.3) relevant. Für die spieltheoretischen Überlegungen ist die Annahme von Absichten und Erwartungen jedoch nicht notwendig, wie wir weiter unten noch sehen werden.

		Fahrer 1	
		links	rechts
Fahrer 2	links	1,1 (sicher)	0,0 (Kollision)
	rechts	0,0 (Kollision)	1,1 (sicher)

Abbildung 36: Vier Spielausgänge des Koordinationsspiels „Brückenüberquerung“. Den jeweiligen Resultaten werden Zahlenwerte (hier 1 und 0) zugeordnet, die die Präferenzen der Spieler wiedergeben. Die erste Zahl steht für die Präferenz des ersten Fahrers, die zweite für die des zweiten Fahrers.

Bei diesem Spiel liegt ein echtes Koordinationsproblem vor, denn es gibt zwei koordinative Gleichgewichte¹⁰³. Wiederholt man das Spiel mit denselben Spielern (Fahrern), dann setzt sich nach einer gewissen Zeit eine der beiden Lösungen (beide fahren links, beide fahren rechts) durch. Die Lösung, die sich durchgesetzt hat, ist insofern arbiträr als es nicht naturnotwendig war, dass sie und nicht die andere Lösung sich durchsetzen würde. Fest stand nur, dass sich eine Lösung durchsetzen würde. Welche es ist, wird durch einen kontingenten Prozess bestimmt. Im Fall der Brückenüberquerung spielen für das Durchsetzen einer der beiden Lösungen natürlich auch die Erwartungen der Fahrer eine wichtige Rolle¹⁰⁴, aber dies ist nicht zwingend für die Lösung von Koordinationsproblemen, wie wir gleich sehen werden.

Verbinden wir zunächst unsere Überlegungen zu den relationalen Eigenfunktionen mit unseren Überlegungen zu den „in Reihe geschalteten“ Eigenfunktionen, dann sehen wir, dass auch hier Koordinationsprobleme auftreten können¹⁰⁵. Angenommen es ist sowohl für die Vorrichtung A als auch die Vorrichtung B von Vorteil, wenn in der Situation 1 (S1) der Effekt 1 (E1) und Situation 2 (S2) der Effekt 2 (E2) produziert wird. Zudem nehmen wir an, dass es jeweils zwei Typen der Vorrichtungen A und B gibt:

A I : produziert P1¹⁰⁶ in S1 und P2 in S2

A II : produziert P2 in S1 und P1 in S2

B I : produziert E1 bei P1 und E2 bei P2

B II : produziert E2 bei P1 und E1 bei P2

¹⁰³ Definition von Lewis: „Ich möchte ein *koordinatives Gleichgewicht* definieren als ein Gleichgewicht, bei dem kein Teilnehmer besser abgeschnitten hätte, wenn *irgendein* Teilnehmer allein anders gehandelt hätte, entweder er selbst oder irgendein anderer.“ (Lewis 1975, 14) Für Lewis gehören zu einem Koordinationsproblem wesentlich mehrere koordinative Gleichgewichte (Vgl. Lewis 1975, 16).

¹⁰⁴ Lewis erklärt das Durchsetzen einer Lösung durch Präzedenzfälle. D.h. eine zunächst zufällig gelungene Lösung des Koordinationsproblems wird zum Präzedenzfall und etabliert sich dadurch. Hierbei spielen nach Lewis die Absichten und Erwartungen der Akteure eine wichtige Rolle. „Die Wirkung des Präzedenzfalls lässt sich genauso erklären, wie wir die Wirkung der Ausgefallenheit erklärt haben. [...] Das Gleichgewicht hat hier seine auffällige Besonderheit darin, daß wir es schon beim vorigen Mal erreicht haben. Wir könnten geneigt sein, die frühere Handlung einfach zu wiederholen, weil wir keinen wichtigen Grund für ein anderes Verhalten sehen. Ob nun einer von uns diese Neigung hat oder nicht – wir können jedenfalls erwarten, daß der andere sie hat, oder erwarten, daß der andere von uns erwartet, daß wir sie haben usw. – d.h. wir haben jeder Erwartungen erster oder höherer Ordnung, daß die andern ihr Teil für das Zustandekommen des früheren koordinativen Gleichgewichts beitragen werden, wenn sie keine besonderen Gründe dagegen haben.“ (Lewis 1975, 37).

¹⁰⁵ Das folgende Beispiel orientiert sich an einem Beispiel von Skyrms (2010, 9-11).

¹⁰⁶ P soll hier für „Produkt“ stehen. Wir können uns z.B. einen Botenstoff oder neutraler einfach ein Protein darunter vorstellen.

Auch hier gibt es zwei koordinative Gleichgewichte, nämlich die Kombination von A I mit B I und die Kombination von A II mit B II.

		Vorrichtung A	
		A I	A II
Vorrichtung B	B I	1,1	0,0
	B II	0,0	1,1

Abbildung 37: Mögliche Kombinationen der Vorrichtungstypen und deren Präferenzwerte¹⁰⁷

Wenn wir zudem annehmen, dass es sich bei den Vertretern der Vorrichtungen A und B um zwei verschiedene Populationen handelt, dann können wir theoretisch zeigen, dass sich eine der beiden Kombinationen durchsetzen wird (Vgl. für die folgende Darstellung Skyrms 2010, 9-11). Dazu betrachten wir die zeitliche Entwicklung der Populationen im Sinne des Modells der *replicator dynamics*¹⁰⁸. Auf diese Weise erhalten wir eine Erklärung für die Stabilisierung einer der beiden Kombinationen, die unseren naturalistischen Ansprüchen genügt.

Die Veränderungen der Populationsanteile berechnen wir dann nach folgender Formel (Vgl. Skyrms 2010, 10):

$$dp(A)/dt = p(A) [U(A) - U]$$

In der Formel ist $U(A)$ die durchschnittliche Auszahlung für Strategie A und U die durchschnittliche Auszahlung in der gesamten Population. In unserem Beispiel gibt es vier Strategien A I, A II, B I und B II. Die Auszahlung der jeweiligen Strategie entspricht dem Anteil der jeweils passenden anderen Strategie. Die Auszahlung für Strategie A I entspricht also dem Populationsanteil von B I, damit ist $U(A I) = p(B I)$. Aus der Formel, insbesondere dem Teil in der eckigen Klammer, geht hervor, dass der Anteil einer Strategie in der Population ansteigen wird, wenn deren durchschnittliche Auszahlung höherer ist als die durchschnittliche Auszahlung in der Population. Die Dynamik der Populationsanteile lässt sich dann mit Hilfe von folgendem Diagramm darstellen.

¹⁰⁷ Präferenz im Sinne von "erhöht die Überlebenschancen".

¹⁰⁸ Vgl. z.B. Hofbauer und Sigmund (1998).

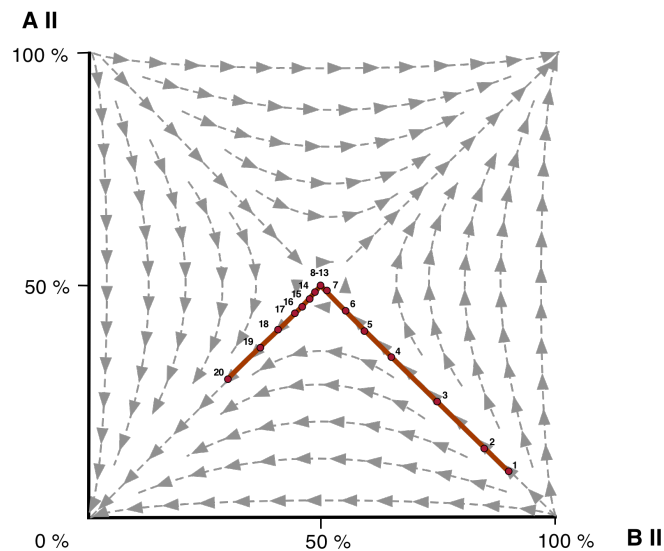


Abbildung 38: Replikatoren Dynamik von A II und B II (in Anlehnung an Skyrms 2010, 11)

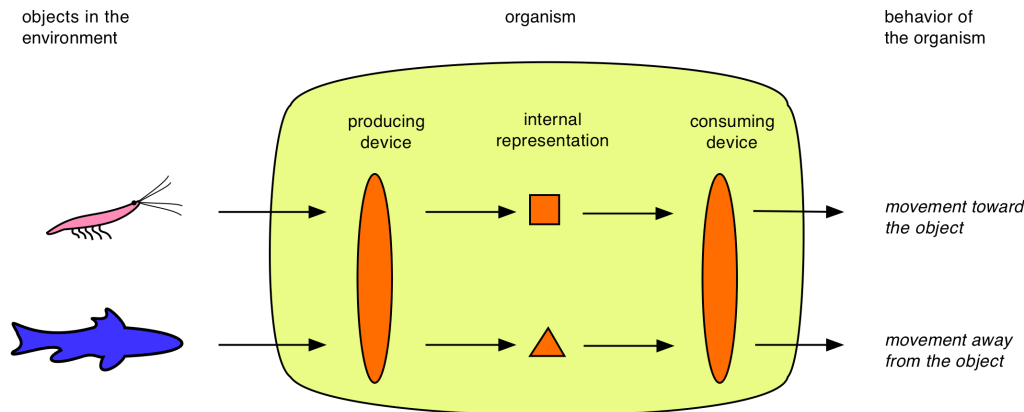
Die Rote Linie im Diagramm zeigt eine mögliche Entwicklung, die auf einer Simulation basiert¹⁰⁹. Der Verlauf ist in seiner Anfangsbewegung zum Zentrum typisch, danach besteht eine 50:50 Wahrscheinlichkeit für eine Bewegung hin zu einem der beiden koordinativen Gleichgewichte. In der Simulation zur Kombination A I/B I (linke untere Ecke), d.h., die Anteile von A II und B II verschwinden am Ende völlig (0 %). Welches der beiden koordinativen Gleichgewichte sich bei einer Simulation tatsächlich einstellen wird, ist also prinzipiell unvorhersagbar. Aber wenn sich eine der beiden Kombinationen einmal durchgesetzt hat, dann ist sie stabil. In diesem Sinne kann man sagen, dass es bei dem simulierten Verlauf zum *Standard* werden wird, in S1 P1 zu produzieren und bei P1 E1 hervorzubringen¹¹⁰. Der Zusammenhang S1-P1-E1¹¹¹ ist insofern arbiträr als es zu ihm eine gleichwertige Alternative (S1-P2-E1) gegeben hat, die sich aber in einem kontingenten Prozess nicht hat durchsetzen können. Mit anderen Worten: Der Zusammenhang S1-P1-E1 besteht nicht aufgrund einer Naturnotwendigkeit.

Betrachten wir jetzt noch einmal das Schema aus Abschnitt 2.1.1.1 mit dessen Hilfe wir uns eine erste Vorstellung von Millikans Zeichenbegriff verschafft hatten.

¹⁰⁹ Für die Simulation wurde eine Anfangsverteilung der Strategien von A I: 90%, A II: 10 %; B I: 10 %, B II: 90% angenommen. Die Simulation lief 20 Runden und es wurde eine Schwankung eingebaut. Für die Simulation wurde folgende Formel verwendet: $dp(A)/dt = p(a) \cdot (U(A) - U) + ((-1)^X) \cdot 0,0001$, wobei $0,0001 \cdot (-1)^X$ die Schwankung ist und X die normalverteilte Zufallsgröße darstellt, also entweder eine 0 oder eine 1, wie bei einem Münzwurf. Ich danke Michael Friedrich für das Programmieren der Simulation und die hilfreichen Diskussionen.

¹¹⁰ Entsprechend: in S2 P2 und bei P2 E2.

¹¹¹ und S2-P2-E2.



Wir hatten dazu Folgendes angenommen:

Für einen im Wasser lebenden Organismus sind Fische Fressfeinde und Krebstiere Nahrung. Von Fischen schwimmt der Organismus weg und auf Krebstiere zu. Zu dieser (kognitiven) Leistung ist der Organismus in der Lage, weil er über ein repräsentationales System verfügt, das als Sender-Empfänger-Modell verstanden werden kann. Der Sender des repräsentationalen Systems (Sender und Empfänger zusammen) ist das sensorische System des Organismus (producing device). Der producing device produziert in Abhängigkeit von bestimmten Umweltbedingungen, Fische oder Krebstiere anwesend, etwas, einen bestimmten Botenstoff oder einen bestimmten Nervenimpuls:

- Wenn Krebstiere anwesend, dann Nervenimpuls A (durch Viereck symbolisiert)
- Wenn Fische anwesend, dann Nervenimpuls B (durch Dreieck symbolisiert)

Der Empfänger des repräsentationalen Systems ist das effektorische System des Organismus (consuming device). Der consuming device bringt in Abhängigkeit von dem, was der producing device produziert hat, ein bestimmtes Verhalten, hin- oder wegschwimmen, hervor:

- Wenn Nervenimpuls A (durch Viereck symbolisiert), dann hinschwimmen
- Wenn Nervenimpuls B (durch Dreieck symbolisiert), dann wegschwimmen

Wir können jetzt sehen, dass wir das repräsentationale System es als eines der beiden soeben behandelten Gleichgewichte interpretieren können. Hierzu müssen wir lediglich wie folgt zuordnen:

Vorrichtung A I – producing device

Vorrichtung B I – consuming device

Situation 1 – Krebs anwesend

Situation 2 – Fisch anwesend

Produkt 1 – Nervenimplus A

Produkt 2 – Nervenimplus B

Effekt 1 – Hinschwimmen

Effekt 2 - Wegschwimmen¹¹²

¹¹² Die auf der Replikatorendynamik basierende Erklärung für das Durchsetzen einer der beiden Lösungen lässt sich allerdings nicht ohne Weiteres auf das Schema übertragen. Hierzu müsste man nämlich annehmen, dass die innerorganismischen Vorrichtungen A und B zu separaten Populationen gehören, was wenig Sinn ergibt. Wir könnten uns aber vorstellen, dass es eine Population von

Insgesamt ergibt sich also Millikans Auffassung eines Zeichens aus der Kombination von spezifischen Eigenfunktionen, nämlich relationalen Eigenfunktionen und Eigenfunktionen, die „in Reihe geschaltet sind“.

Organismen gibt, die Vorrichtungen vom Typ A I und B II beherrbergen. Durch Mutation von A I zu A II oder von B II zu B I würden dann Individuen mit einem Selektionvorteil entstehen. Welche der beiden Kombinationen (A I/B I oder A II/B II) realisiert werden würde, hinge wahrscheinlich davon ab, welche der beiden Mutationen (A I zu A II oder B II zu B I) zuerst entsteht. Aber auch hier würde sich diejenige Kombination (A I/B I oder A II/B II), die zuerst entsteht, nach ihrer Ausbreitung als stabil erweisen. In diesem evolutionstheoretischen Sinne sollte man auch die Passage in Millikans Zeichendefinition interpretieren, in der davon Rede ist, dass der *producing device* und *consuming device*, so designt sind, dass sie zueinander passen (“designed [...] to fit one another”). Die Passage gehört zur zweiten Bedingung von Millikans Zeichendefinition. Zur Erinnerung die Passage lautet: “Normally a sentence stands midway between two cooperating devices, a producer device and an interpreter device, which are designed or standardized to fit one another, the presence and cooperation of each being a Normal condition for the proper performance of the other.” (Millikan 1984, 97).

2.1.1.3.8 Mapping Function

In diesem Abschnitt soll erläutert werden, was Millikan unter einer mapping function (Abbildungsfunktion) versteht.

Um einem Missverständnis vorzubeugen: Mapping functions sind keine spezifischen Eigenfunktionen, sondern Funktionen im mathematischen Sinn. Die Grundidee bei einer mapping function ist die von einer eindeutigen Zuordnung von Elementen aus einem Bereich (domain) zu Elementen aus einem weiteren Bereich (range)¹¹³. Eindeutig ist die Zuordnung, wenn jedem Element der domain genau ein Element der range zugeordnet werden kann.

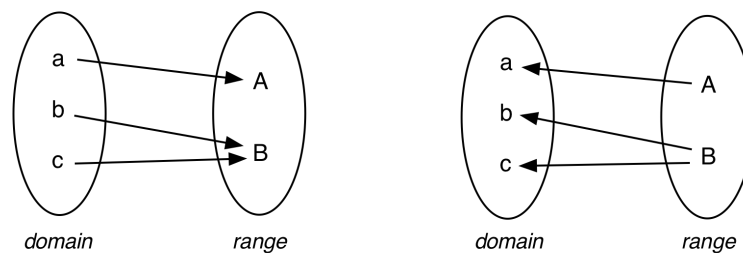


Abbildung 39: Eindeutige und nicht eindeutige Zuordnungen

In der Abbildung 39 ist also die erste Zuordnung eindeutig, die zweite hingegen nicht. Ist eine wechselseitige oder eineindeutige Zuordnung zwischen zwei Bereichen möglich (siehe Abbildung 40), dann bezeichnet man diese als 1:1 Korrespondenz oder auch als Isomorphismus.

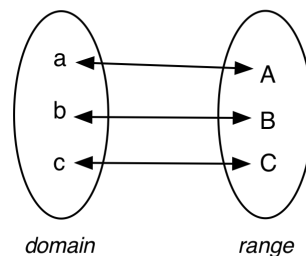


Abbildung 40: Eineindeutige Zuordnung

Sobald die Anzahl der Elemente in den beiden Bereichen gleich ist, lässt sich immer eine solche Zuordnung bilden, die meist weitere Alternativen hat.

¹¹³ Meine Darstellung der mapping function orientiert sich an Shea (2013, 64-65, 74-75).

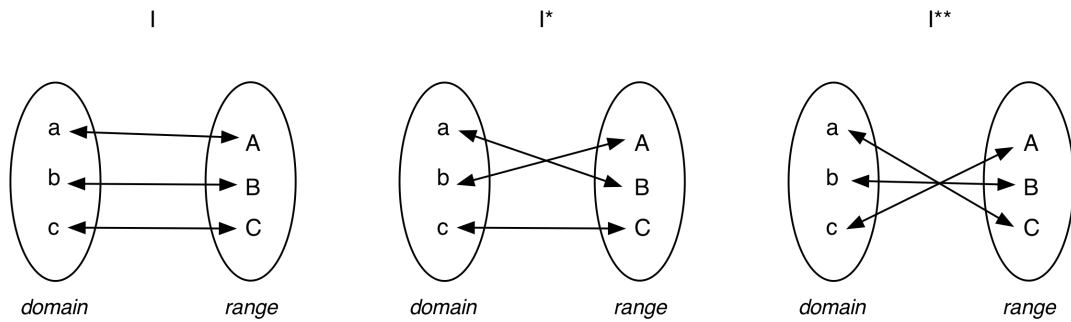
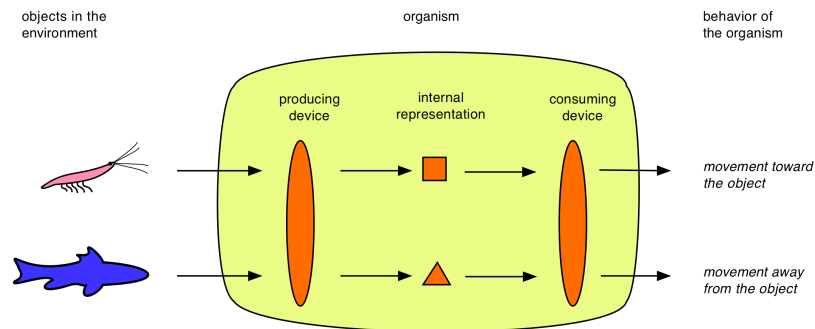


Abbildung 41: Eineindeutige Zuordnung (I) mit Alternativen (I*, I**)

Die Behauptung, dass zwischen zwei Bereichen ein Isomorphismus besteht, ist also nicht besonders stark. So besteht zwischen den acht Stiften in meiner Tasche und den Planeten im Sonnensystem bei geeigneter Zuordnung eine Isomorphie.

Betrachten wir nun wieder unser Schema vom Anfang, das das repräsentationale System eines fiktiven Wasserlebewesens darstellt.



Auch hier bestehen Isomorphismen (siehe Abbildung 42). Den Umweltzuständen (Nahrung anwesend, Fressfeind anwesend) sind eineindeutig Zwischenprodukte (Viereck, Dreieck) zugeordnet und diesen wiederum Verhaltensäußerungen (Annähern, Fliehen), so dass letztlich auch eine Isomorphie zwischen den Umweltzuständen und den Verhaltensäußerungen besteht.



Abbildung 42: Isomorphismen bei in unserem fiktiven Beispiel

D.h. die Normale Erklärung, wie z.B. das durch ein Dreieck symbolisierte Zwischenprodukt den consuming device anpasst (adapts), so dass dieser seine Eigenfunktion ausüben kann, muss erwähnen, dass dem durch ein Dreieck symbolisierten Zwischenprodukt der Umweltzustand „Fressfeind anwesend“ zugeordnet ist.

Bevor wir im nächsten Abschnitt klären, welche Rolle die genannten Isomorphismen bei der Gehaltsbestimmung spielen, wollen wir noch kurz auf die Bedeutung der mapping function für die Produktivität eines Repräsentationssystems eingehen.

Die Produktivität eines Zeichensystems wird häufig als Voraussetzung dafür angesehen, dass auch neue Zeichen(ketten) verstanden werden können¹¹⁴. Da Millikan beansprucht, dass ihre Theorie auch auf die menschliche Sprache angewendet werden kann, muss sie dem Verstehen neuer Zeichen auch in ihrer Theorie Rechnung tragen. Dies hat gelegentlich den Einwand provoziert, dass eine Theorie, die sich so stark auf Selektionsgeschichte stützt, hierbei versagen muss (siehe auch S. 83). Der Einwand ist aber unberechtigt.

Die folgenden Überlegungen können als Exkurs betrachtet werden, da der Einwand von Millikans Anspruch (Vgl. Abschnitt 1.4.3) ausgeht, der in dieser Arbeit nicht vertreten wird.

Wir können erkennen, dass der Einwand unberechtigt ist, wenn wir die mapping function nicht wie bisher als Zuordnungsliste verstehen, sondern als allgemeiner gefasste Zuordnungsregel. Besonders deutlich wird dies an Millikans Lieblingsbeispiel dem Schwänzeltanz der Bienen.

„Under certain circumstances, the foreaging bee returning to the hive from a desirable resource performs a waggle dance inside the hive. [...] To a first approximation, the dancing honeybee repeatedly traverses a figure-8 pattern whose two circles meet in a straight line. While traversing the straight line, the bee repeatedly waggles its abdomen. The average orientation of the bee's waggle run with respect to gravity reliably correlates with the food source's solar bearing [...]. The average duration of the waggle run reliably correlates with distance from the hive to the food source. The dance recruits various bees, who promptly fly toward the vicinity of the food source.“ (Rescorla 2013, 90f.)

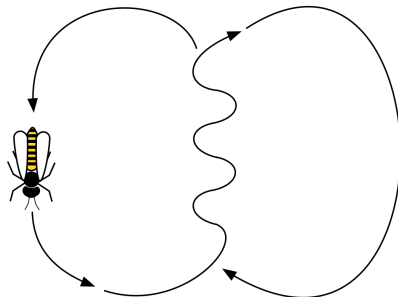

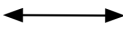
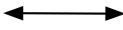
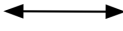
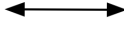


Abbildung 43: Schwänzeltanz (waggle dance) der Honigbiene

Betrachten wir nun folgende Zuordnung von Schwänzelsbewegungen und Entfernungen¹¹⁵:

¹¹⁴ „*Argument from productivity*: Since competent speakers can understand a complex expression *e* they never encountered before, it must be that they (perhaps tacitly) know something on the basis of which they can figure out, without any additional information, what *e* means. If this is so, something they already know must determine what *e* means. And this knowledge cannot plausibly be anything but knowledge of the structure of *e* and knowledge of the individual meanings of the simple constituents of *e*.“ (Szabó 2013), (Szabó, Zoltán Gendler, "Compositionality", *The Stanford Encyclopedia of Philosophy* (Fall 2013 Edition), Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/fall2013/entries/compositionality/>>.)

¹¹⁵ Das Beispiel ist von Shea (2013, 74) übernommen.

1 waggle		300m
2 waggles		150m
3 waggles		100m
4 waggles		75m
5 waggles		60m

Wir können die Zuordnung wie zuvor als Liste auffassen. Wir können sie aber auch als allgemeinen Zusammenhang verstehen, der durch folgende Regel wiedergegeben werden kann: *Anzahl x der Schwänzelsbewegungen* wird der Entfernung $300m/Anzahl\ x\ der\ Schwänzelsbewegungen$ zugeordnet.

Durch die allgemeine Regel sind nun auch Zuordnungen möglich, die nicht in der Liste erfasst sind. Eine sehr nahe Nektarquelle in der Entfernung von 50m, würde z.B. einen Tanz mit 6 Schwänzelsbewegungen zur Folge haben. Theoretisch wäre es so auch möglich, wenn auch praktisch unwahrscheinlich, dass noch nie in der Bienengeschichte eine Nektarquelle 150m vom Stock entfernt war. Verhalten sich die Bienen aber entsprechend der oben genannten Regel, wird die von der 150m entfernten Nektarquelle heimkehrende Biene trotzdem einen Tanz mit 2 Schwänzelsbewegungen aufführen.

Eine als allgemeine Regel verstandene mapping function erlaubt es Millikan also prinzipiell auch neue Zeichen in ihrer Theorie zu erfassen.

2.1.2 Gehaltsbestimmung bei Millikan

In diesem Abschnitt soll erläutert werden, wodurch in Millikans Theorie der Gehalt eines Zeichens bestimmt ist.

Die Gehaltsbestimmung bei Millikan ist relativ komplex und auch nicht ganz eindeutig. Millikan liefert nirgends eine Definition in Sinne von „Der Gehalt eines Zeichens y ist ...“ und unterschiedliche Interpretationen der Gehaltsbestimmung sind aufgrund einiger Textstellen möglich. Ich werde später in diesem Abschnitt noch auf Interpretationen eingehen, die Millikans Position unnötig schwächen und die ich daher ablehne. Einige unterschätzten die Rolle der Abbildungsfunktion bei der Gehaltsbestimmung, während die andere sie überschätzten.

Insgesamt unterscheide ich hier insgesamt vier Interpretationen:

1. meine eigene Interpretation (kurz I 1)
2. eine Interpretation, die die Abbildungsfunktion unberücksichtigt lässt (kurz I 2)
3. eine zweite Interpretation, die die Abbildungsfunktion unberücksichtigt lässt (kurz I 3)
4. eine Interpretation, die die Bedeutung der Abbildungsfunktion überschätzt (kurz I 4)

Millikan veranschaulicht ihre Ausführungen zur Gehaltsbestimmung oft mit Beispielen, was an sich lobenswert ist. Es führt aber m.E. dazu, dass man eher intuitiv erfasst, wie bei Millikan der Gehalt bestimmt ist, anstatt dies mit Hilfe der von der Theorie bereitgestellten Werkzeuge nachzuvollziehen.

Ich werde daher im Folgenden die Gehaltsbestimmung bei Millikan am Beispiel des uns nun bereits vertrauten Wasserlebewesens *unter expliziter Berücksichtigung der theoretischen Mittel ihrer Theorie erläutern*.

Erläuterung an fiktivem Beispiel

Zunächst ist daran zu erinnern, dass es sich bei den inneren Zeichen unseres Wasserlebewesens um Pushmi-Pullyu-Repräsentationen handelt, also um Zeichen, die sowohl einen indikativen als auch einen imperativen Gehalt haben (Vgl. S. 49). Ich konzentriere mich im Folgenden aber nur auf den indikativen Gehalt.

Angenommen, uns ist die grundlegende Struktur des repräsentationalen Systems (Producer, Icon, Consumer) der Wasserlebewesen bekannt. Bei einer weiteren Untersuchung eines der Wasserlebewesen stellen wir dann fest, dass unter Bedingung x das „Dreiecks-Icon“ produziert wurde und dieses wiederum Verhalten y ausgelöst hat (siehe Abbildung 44).

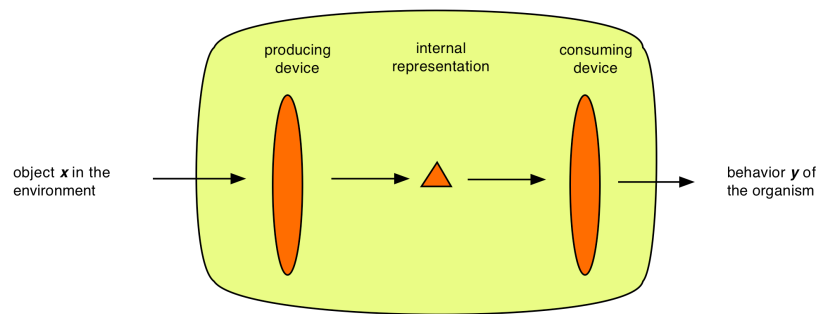


Abbildung 44: Ursache und Wirkung des „Dreiecks-Icons“

Was ist nun der indikative Gehalt des „Dreiecks-Icons“? Um dies zu bestimmen, müssen wir etwas anderes kennen als den beobachteten kausalen Zusammenhang, der für die Gehaltsbestimmung nicht ausreichend ist. Glücklicherweise – so nehmen wir an – sind unsere Wasserlebewesen im Labor entstanden. Die gesamten Bedingungen ihrer Evolution wurden kontrolliert und dokumentiert. Ebenso ist uns die Physiologie der Lebewesen bestens bekannt. D.h., wir wissen, dass der Producer entweder „Dreieck-Icons“ oder „Viereck-Icons“ produzieren kann. Wir wissen auch, dass es eine Phase in der Evolution dieser Lebewesen gegeben hat, in der sich die Kombination (Verschaltung) *Schrimp-Viereck-hin; Fisch-Dreieck-weg* gegenüber der Kombination *Schrimp-Dreieck-hin; Fisch-Viereck-weg* durchgesetzt hat (siehe Abbildung 45).

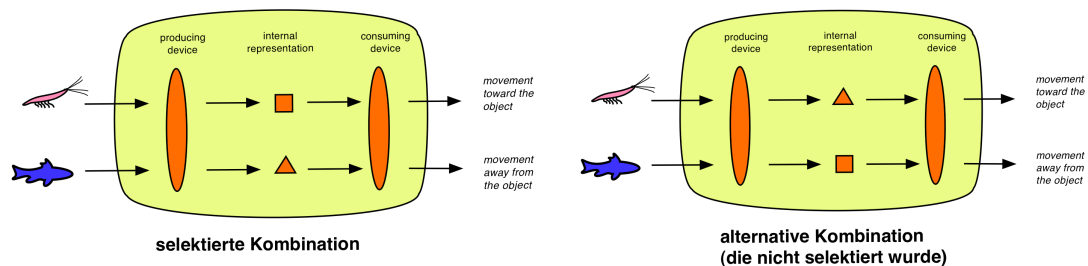


Abbildung 45: Selektierte und alternative (nicht selektierte) Kombination von internen Repräsentationen und In- und Output

Wir wissen also, dass die Icons eine *standardisierende Funktion* in Millikans Sinne haben (Vgl. Abschnitt 2.1.1.3.7).

Obwohl die Lebewesen rein physiologisch gesehen auf kleine bzw. große dunkle Objekte reagieren und nicht nur auf Schrimps und Fische, ist die oben genannte Zuordnung in Millikans Sinne korrekt, denn es waren nur die Begegnungen mit Schrimps und Fischen, bei denen das durch das repräsentationale System gesteuerte Verhalten einen Selektionsvorteil bewirkt hat. Dies – so nehmen wir wieder großzügig an – kann durch genaue Aufzeichnungen und Messungen belegt werden¹¹⁶. Wir wissen also, dass Anwesenheit von Schrimps bzw. Fischen die *Normalen Bedingungen* (Vgl. Abschnitt 2.1.1.3.4) für das Ausüben der *relationalen Eigenfunktionen* (Vgl. Abschnitt 2.1.1.3.6) von Producer und Consumer sind.

Für die Gehaltsbestimmung des Dreieck-Icons ist nun entscheidend, dass wir uns auf dessen *direkte Eigenfunktion* und nicht auf dessen *abgeleitete Eigenfunktion* (Vgl. Abschnitt 2.1.1.3.6) konzentrieren. Eine sehr wichtige und m.E. viel zu wenig beachtete Passage aus Millikan (1984) macht dies deutlich.

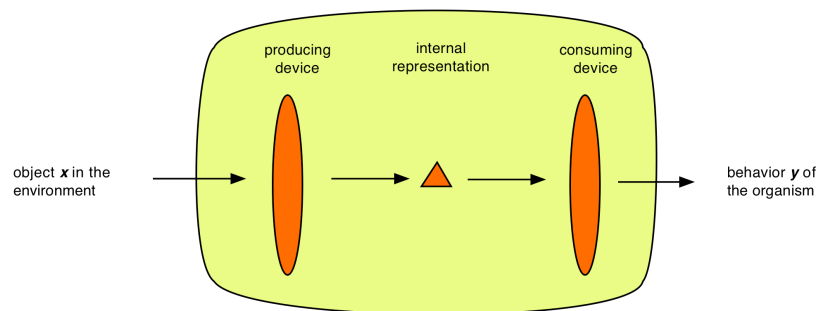
¹¹⁶ Wir können uns dies etwa analog zu den Studien von Kettlewell (1955) zum Industriemelanismus vorstellen (Vgl. S. 62).

„In Part II I will argue that the most dominant notion of what is signed by signs is derived by reference to direct proper functions of these signs themselves, hence to resulting adapted proper functions of interpreting devices qua taking these signs as immediate adaptors. It is not derived by reference to adapted functions of the sign's producing devices. (On the other hand, sometimes we do consider a sign to be "about" its original adaptor.)“ (Millikan 1984, 43)

In unserem Beispiel ist es die direkte relationale Eigenfunktion des Icons den Consumer an die der eigenen Form zugeordneten Umweltbedingung anzupassen. Im konkreten vorliegenden Fall (Abbildung 44) hat das Dreiecks-Icon also die direkte Eigenfunktion den Consumer an die der Dreiecksform zugeordnete Umweltbedingung: Fisch anwesend anzupassen. Der Gehalt des konkret vorliegenden Dreiecks-Icons nach meiner Interpretation ist also: Fisch anwesend.

Die abgeleitete Eigenfunktion des Dreieck-Icons, kann mit dessen direkter Eigenfunktion überstimmen, aber sie muss es nicht. Die abgeleitete Eigenfunktion des Dreieck-Icons ist es den Consumer an die Umweltbedingung anzupassen, an die der Producer laut seiner Eigenfunktion das Icon hätte anpassen müssen.

Wenden wir uns zur Veranschaulichung nochmals der Abbildung 44 zu.



Die direkte Eigenfunktion des Dreieck-Icons steht fest. Sie besteht darin den Consumer an die der Dreiecksform zugeordnete Umweltbedingung „Fisch anwesend“ anzupassen. Die abgeleitete Eigenfunktion des Dreieck-Icons hingegen hängt davon ab, um welche Umweltbedingung es sich bei der Bedingung x handelt. Ist es „Fisch anwesend“, dann entspricht die abgeleitete Eigenfunktion des Dreieck-Icons seiner direkten Eigenfunktion. Ist es aber „Schrimps anwesend“¹¹⁷, dann besteht ein Konflikt zwischen der abgeleiteten Eigenfunktion und der direkten Eigenfunktion des Dreieck-Icons, denn die abgeleitete Eigenfunktion wäre es den Consumer an „Schrimps anwesend“ anzupassen.

Der Fall „Schrimps anwesend“ und Produktion des Dreieck-Icons wäre eine Missrepräsentation, wie wir sie in Abschnitt 1.3.3 beschrieben haben. Würde hingegen die abgeleitete Eigenfunktion des Icons den Gehalt des Icons bestimmen, dann wäre eine Missrepräsentation nicht möglich. Warum dies so ist wird klar werden, wenn wir unten auf die Interpretation I 2 eingehen. Wir werden aber später noch untersuchen müssen, wie Millikan das Eintreten eines solchen Falles („Schrimps anwesend“ und Produktion des Dreieck-Icons) erklärt (Abschnitt 2.2) und ob ein solcher Fall unter der Annahme von Millikans Theorie überhaupt eintreten könnte (Teil 3). D.h. in Teil 3 wird es um die Frage gehen, ob Millikan das, was wir in einem solchen Fall intuitiv als „schiefgelaufen“ betrachten auch theoretisch als ein Schiefgelaufen erklären kann, und zwar ohne nicht-naturalistische Anleihen aufzunehmen.

¹¹⁷ Der Einfachheit halber ignoriere ich hier die Fälle: „weder Fisch noch Schrimps anwesend“ und „beide, Fisch und Schrimps, anwesend“.

Zunächst können wir aber festhalten, dass der Gehalt des Dreieck-Icons nach meiner Millikan Interpretation (I 1) also „Fisch anwesend“ ist und zwar unabhängig davon, ob tatsächlich Fisch anwesend ist.

Allgemein könnte man den Gehalt eines Icons nach Millikans Theorie vielleicht, wie folgt definieren¹¹⁸:

Gehalt des Icons = diejenige Umweltbedingung (Normale Bedingung), die vorliegen müsste, damit der Consumer seine (adaptierte) Eigenfunktion erfüllen kann, wenn er das Icon als unmittelbaren Adaptor nimmt.

In unserem Beispiel gehört es zur (adaptierten) Eigenfunktion des Consumers auf das Dreiecks-Icon mit einer Fluchtbewegung zu reagieren, der Consumer kann aber seine so (adaptierte) Eigenfunktion nur dann erfüllen, wenn auch tatsächlich ein Fressfeind (Fisch) anwesend ist. Die Anwesenheit eines Fressfeindes (Fisches) ist also die Normale Bedingung für die (unmittelbar an das Icon adaptierte) Eigenfunktion des Consumers und damit der Gehalt des Icons.

Ohne direkt auf die Funktion des Consumers einzugehen, könnte man auch sagen, dass der Gehalt eines Icons, diejenige Bedingung ist, die Normalerweise mit der spezifischen Form des Icons korrespondiert.

"The content hangs only on there being a certain condition that would be normal for performance of the consumer's functions – namely, that a certain correspondence relation hold between sign and world – whatever those functions may happen to be." (Millikan 1989, 287)

"[...] represented conditions are conditions that vary, depending on the form of the representation, in accordance with specifiable correspondence rules that give the semantics for the relevant system of representation." (Millikan 1989, 287)

Dass allerdings eine solche Korrespondenz zwischen Umweltbedingungen und Zeichenformen besteht, beruht darauf, dass zwei miteinander kooperierende Mechanismen selektiert wurden, einer, der beim Vorliegen bestimmter Umweltbedingungen bestimmte Zeichenformen hervorbringt und einer, der auf diese Zeichenformen hin spezifische Verhaltenäußerungen bewirkt (Vgl. Abschnitt 2.1.1.3.7).

Insgesamt können wir zudem festhalten, dass eine korrekte Gehaltsbestimmung eines Icons bei Millikan ein recht umfangreiches Wissen über die Evolution der Iconträger/-nutzer erfordert. Ein solches Wissen wird uns nur in den allerwenigsten Fällen vorliegen, so dass die Gehaltsbestimmung nach Millikan fast immer spekulative Züge trägt. In Hinsicht auf den ambitionierten Anspruch von Millikan, auch die menschliche Sprache und Gedanken mit ihrer Theorie zu erfassen, könnte man nun einwenden, dass der Gehalt sprachlicher Ausdrücke und Gedanken faktisch keinen solchen spekulativen Charakter hat. Wir wissen, was wir sagen und denken. D.h., wir kennen den Gehalt unserer Gedanken und Äußerungen. Genau dies bestreitet Millikan konsequenterweise. Ob zu Recht oder zu Unrecht, ist aber nicht Gegenstand der vorliegenden Untersuchung (Vgl. auch Abschnitt 1.4.3).

Ich werde nun auf Interpretationen der Gehaltsbestimmung bei Millikan eingehen, die Millikans Position unnötig schwächen und die ich daher ablehne. Die einen unterschätzten

¹¹⁸ Es finden sich erstaunlich wenige Versuche einer solchen Definition in der Sekundärliteratur zu Millikan. Eine Ausnahme bildet Stegman (2009). "In a nutshell, the (descriptive) content of a signal is that circumstance in the world the mapping of which (to signal tokens of this type) is (a) the function of the producing device and (b) the normal condition for the consuming device to achieve its function (Millikan 2004, 76–80)." (Stegmann 2009, 867).

die Rolle der Abbildungsfunktion bei der Gehaltsbestimmung (I2 und I3), während die andere sie überschätzt (I4).

Beginnen wir mit der Interpretation I 2, die die Rolle der Abbildungsfunktion unterschätzt. Nach dieser Interpretation sind einzig und allein der Consumer und dessen Eigenfunktion für die Gehaltsbestimmung relevant.

Gehalt des Icons = diejenige Umweltbedingung (Normale Bedingung), die vorliegen muss, damit der Consumer des Icons seine Eigenfunktion erfüllen kann.

Dies läuft darauf hinaus, dass die abgeleitete Eigenfunktion und nicht die direkte Eigenfunktion des Icons den Gehalt bestimmt. Die abgeleitete Eigenfunktion des Icons ist es den Consumer an die gerade vorliegende Umweltsituation anzupassen.

Millikans starke Betonung des Consumers und Formulierungen, wie die folgende, scheinen diese Interpretation zu stützen:

"Indeed, a good look at the consumer part of the system ought to be all that is needed to determine not only representational status but representational content." (Millikan 1989, 286)

Ich werde im Teil 3 auf eine konkrete Kritik eingehen, der diese Interpretation zugrunde zu liegen scheint (Vgl. Abschnitt 3.2.6). Es lassen sich aber in der Sekundärliteratur häufig Formulierungen finden, die diese Interpretation wiedergeben. So schreibt z.B. Markus Wild:

„Schließlich ist die Biosemantik eine konsumentenorientierte Version der Teleosemantik. Dies hat insbesondere Auswirkungen auf die Wahrnehmungstheorie. *Wird der Inhalt von Wahrnehmungen durch die Bedingung festgelegt, die vorliegen muss, damit ein Konsument seine Funktion ausüben kann?* [Hervorh. AW]" (Wild 2010, 15)¹¹⁹

Die Interpretation ist sicherlich schon deshalb attraktiv, weil sie sich die komplizierten Ausführungen von Millikan zur Abbildungsfunktion sparen kann. Zudem scheint sie auch bei der Anwendung auf Millikans Beispiele das richtige Resultat zu liefern.

Ein Bieber¹²⁰ (Produzent) erzeugt beim Anblick eines Fressfeindes einen Kellenschlag (Icon)¹²¹, woraufhin andere Bieber (Konsumenten) Sicherheit im Wasser suchen. Laut Millikan ist der Gehalt des Kellenschlags (Icon) „Gefahr“ (Vgl. Millikan 1989, 288) und zu diesem Ergebnis kommt man auch nach der gerade behandelten Interpretation: Die Anwesenheit einer Gefahrenquelle (der Fressfeind) ist diejenige Umweltbedingung, die vorliegen muss, damit die Fluchtreaktion der Bieber nicht umsonst war, sprich die Consumers ihre Eigenfunktion erfüllen können.

Die Interpretation ist dennoch unhaltbar, denn sie macht Missrepräsentationen unmöglich, was absolut nicht im Sinne Millikans sein kann. Wir können dies zeigen, indem wir uns wieder unserem fiktiven Wasserlebewesen zuwenden. Betrachten wir nochmals die Situation, die nach meiner Interpretation eine Missrepräsentation darstellte. In dieser Situation war Schrimps anwesend und dennoch wurde ein Dreiecks-Icon produziert (siehe Abbildung 46).

¹¹⁹ Wild vertritt die genannte Interpretation jedoch nicht, sondern bedient sich hier nur einer verkürzten Formulierung.

¹²⁰ Genaugenommen sind nicht die Bieber als Ganze Produzenten und Konsumenten, sondern wiederum Mechanismen in ihnen. Vgl. hierzu auch Wild (2010, 44-45).

¹²¹ D.h. er schlägt mit der sogenannten Bieberkelle, das ist sein abgeflachter Schwanz, auf die Wasseroberfläche.

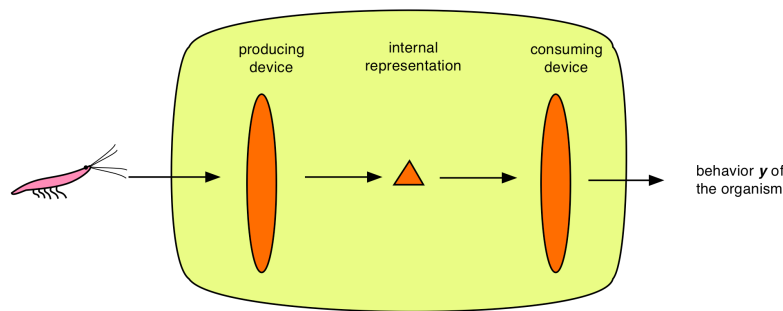


Abbildung 46: Bei Anwesenheit von Schrimps wird Dreiecks-Icon produziert

Nach der gerade behandelten Interpretation liegt hier jedoch keine Missrepräsentation vor. Der Gehalt des Dreiecks-Icons wäre „Schrimps anwesend“, denn dies ist die Bedingung, die vorliegen muss, damit der Consumer in dieser Situation seine (relationale) Eigenfunktion ausüben kann.

Dasselbe Resultat erhalten wir anhand des Bientanzbeispiels. Betrachten wir folgende Situation: Die Nektarquelle ist 300 m entfernt, eine Biene kehrt von dieser Quelle zurück, führt einen bestimmten Tanz auf, woraufhin eine weitere Biene einen bestimmten Flug unternimmt (siehe Abbildung 47).

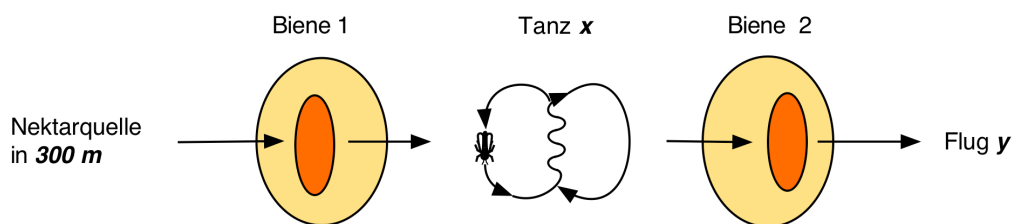


Abbildung 47: Welchen Gehalt hat der Tanz x?

Ohne die bestimmte Form des Tanzes berücksichtigen zu müssen – hierin liegt die Unterschätzung der Abbildungsfunktion für die Gehaltsbestimmung – ist der Gehalt des Tanzes nach der gerade behandelten Interpretation „Nektarquelle in 300m Entfernung“. Dies ist so, weil trivialerweise in einer Situation, in der die Nektarquelle in 300m Entfernung ist, ein an eine 300m entfernte Nektarquelle adaptierter Flug dazu führt, dass die konsumierende Biene¹²² ihre relationale Eigenfunktion erfüllt.

Generell gilt also für diese Interpretation: Da die Bedingung an die der konsumierende Mechanismus angepasst werden muss, damit er seine Eigenfunktion erfüllen kann, immer zugleich diejenige ist, die gerade vorliegt und diejenige, die den Gehalt bestimmt, kann es keine Missrepräsentationen geben. Wenn X vorliegt, muss der Consumer an X angepasst werden und „X“ ist der Gehalt des Icons; wenn Y vorliegt, muss der Consumer an Y angepasst werden und „Y“ ist der Gehalt des Icons; usw.

Eine zweite Interpretation I 3, die die Abbildungsfunktion ignoriert, besteht darin sich auf den tatsächlichen Effekt des Konsumenten zu konzentrieren.

Gehalt des Icons = diejenige Umweltbedingung (Normale Bedingung), die vorliegen müsste, damit der tatsächlich beobachtete Effekt des Consumer einen Selektionsvorteil bringt

¹²² Genauer der Mechanismus in der Biene, die den Tanz beobachtet.

Wir können uns die Gehaltsbestimmung nach I 3 wieder anhand des Bienentanzbeispiels klar machen. Wie zuvor berücksichtigen wir dabei der Einfachheit halber nur die Entfernung der Nektarquelle und nicht die Richtung. Folgende Zuordnung soll wieder gelten (Vgl. Abschnitt 2.1.1.3.8):

1 waggle	↔	300m
2 waggles	↔	150m
3 waggles	↔	100m
4 waggles	↔	75m
5 waggles	↔	60m

Betrachten wir die folgenden vier Fälle:

Fall	Nektarquelle (NQ)	Tanz	Flug	Gehalt nach I 1	Gehalt nach I 3
1	300m	1 waggle	300m ¹²³	NQ in 300m	NQ in 300m
2	300m	2 waggle	150m	NQ in 150m	NQ in 150m
3	150m	2 waggle	300m	NQ in 150m	NQ in 300m
4	300m	2 waggle	300m	NQ in 150m	NQ in 300m

Im ersten Fall ist der Gehalt des Tanzes mit einer Schwänzelsbewegung sowohl nach meiner Interpretation als auch nach I 3: „Nektarquelle 300m entfernt“. Nach meiner Interpretation liegt dies wesentlich an der oben gegebenen Zuordnung. Nach I 3 liegt es daran, dass ein 300m-Flug nur dann einen Vorteil bringt, wenn die Nektarquelle 300m entfernt ist.

Im zweiten Fall haben wir wieder eine Übereinstimmung. Der Gehalt des Tanzes mit zwei Schwänzelsbewegungen ist: „Nektarquelle 150m entfernt“. Nach meiner Interpretation liegt dies an der Zuordnung. Nach I 3 liegt es daran, dass ein 150m-Flug nur dann einen Vorteil bringt, wenn die Nektarquelle 150m entfernt ist. Nach beiden Interpretationen haben wir es hier also mit einer Missrepräsentation zu tun.

Im dritten Fall haben wir keine Übereinstimmung. Nach meiner Interpretation ist der Gehalt des Tanzes mit zwei Schwänzelsbewegungen wieder: „Nektarquelle 150m entfernt“. Nach I 3 hingegen „Nektarquelle 300m entfernt“. Nach meiner Interpretation repräsentiert der Tanz korrekt nach I 3 ist er eine Missrepräsentation.

Im vierten Fall besteht wieder keine Übereinstimmung zwischen den Interpretationen. Nach meiner Interpretation ist der Gehalt des Tanzes mit zwei Schwänzelsbewegungen wieder: „Nektarquelle 150m entfernt“. Nach I 3 hingegen wieder „Nektarquelle 300m entfernt“. Diesmal ist der Tanz aber nach meiner Interpretation eine Missrepräsentation und nach I 3 repräsentiert der Tanz korrekt.

I 3 ist in sofern eine echte Alternative als sie zumindest Missrepräsentationen zulässt. Meines Erachtens wird in dieser Interpretation aber der Unterschied zwischen *ein falsches Zeichen richtig lesen* und *ein richtiges Zeichen falsch lesen* verwischt. Dies führt dann auch dazu, dass nach I 3 kein Unterschied zwischen den Fällen 1 und 4 besteht, obwohl lax gesprochen im Fall 4 einiges schiefgeht. Insgesamt ist aber I 3 keine angemessene Interpretation von Millikans

¹²³ Lies: Die von der 300m entfernten Nektarquelle heimgekehrte Biene führt einen Tanz mit 1 waggle auf, daraufhin unternimmt eine andere Biene einen 300m-Flug (in die entsprechende Richtung).

Position, weil die Abbildungsfunktion völlig unberücksichtigt bleibt. Die Abbildungsfunktion spielt für Millikan eine wichtige Rolle, die teilweise überschätzt wird¹²⁴.

Kommen wir nun zur Interpretation I 4, die die Rolle der Abbildungsfunktion überschätzt. Der Unterschied zwischen I 1 und I 4 liegt nur in der unterschiedlichen Bewertung der explanatorischen Rolle der Abbildungsfunktion. Die Gehaltsbestimmung ist nach I 1 und I 4 gleich.

Nach I 4 kommt der Abbildungsfunktion bei der Gehaltsbestimmung eine analoge Rolle zu, wie derjenigen, die die Indikation bei Dretskes Ansatz spielt¹²⁵ (Vgl. Godfrey-Smith 1996, 185). Dretske geht davon aus, dass es Indikatoren (natürliche Zeichen) unabhängig von irgendwelchen funktionalen Zusammenhängen gibt. Diese können dann durch Selektionsprozesse zu echten Zeichen (Repräsentationen) werden. Analog gilt dann für Millikan, es gibt Dinge, die in einer Isomorphie-Relation (Vgl. Abschnitt 2.1.1.3.8) zu anderen Dingen stehen, und zwar ganz unabhängig von irgendwelchen funktionalen Zusammenhängen. Durch Selektionsprozesse werden diese „Abbildner“ zu echten Zeichen (Repräsentationen).

„Mapping’ would play a similar role to that played by indication in hybrid theories like Dretske’s. Inner states would antecedently map, and be promoted into representations in virtue of selection processes; they would be promoted from mere de facto mappers into things that are supposed to map, just as Dretske has states promoted from de facto indicators into states that are supposed to indicate. Both theories would make use of a proto-semantic relation - indication for Dretske, mapping for Millikan - and would assert that this relation is promoted into a genuinely semantic relation by selection processes.“ (Godfrey-Smith 1996, 185)

Nach dieser Interpretation ist bei Millikan die Isomorphie-Relation die Basis für die Bestimmung des Gehalts und die teleologischen Aspekte dienen nur dazu Missrepräsentationen möglich zu machen. Diese Interpretation findet Rückendeckung durch etliche Textstellen, in denen Millikan explizit diese Position vertritt.

„[...] naturalist theories of the content of mental representation are often divided into, say, picture theories, causal or covariation theories, information theories, functionalist or causal role theories, and teleological theories, as though these divisions all fell on the same plane. That is a fairly serious mistake, for what teleological theories have in common is not any view about the nature of representational content. ‘Teleosemantics,’ as it is sometimes called, is a theory only of how representations can be false or mistaken, which is a different thing entirely. Intentionality, if understood as the property of “ofness” or “aboutness,” is not explained by a teleological theory. [...]“ (Millikan 2004, 63 zitiert nach Shea 2013, 72f)

„[...] What teleological theories do not have in common is an agreed on description of what representing — what “ofness” or “aboutness”— is. They are not agreed on what an organism that is representing things correctly, actually representing things, is doing, hence on what it is that an organism that is misrepresenting is failing to do. To the shell that is ‘teleosemantics’ one must add a description of what actual representing is like.

¹²⁴ Auch von Millikan selbst.

¹²⁵ Markus Wild vertritt m.E. eine solche Interpretation. Wild unterscheidet zwischen R-Inhalten (R für repräsentational) und IR-Inhalten (I für intentional), wobei sich R-Inhalte aus dem Puren bestehen einer Abbildungsrelation ergeben, während IR-Inhalt erst mit der Verwendung durch einen Konsumer entstehen (Vgl. Wild 2013, 34-49). So schreibt Wild beispielsweise in Bezug auf die Dretskes Magnetosome (Vgl. Abschnitt 1.3.4.2): "Das Magnetosom bezeichnet alle diese Dinge [magnetic north, geomagnetic north, deeper water, or lesser oxygen, AW] als R-Inhalte, denn die Orientierung der Magnetosome unterhält Abbildungsrelationen zu jedem der vier genannten Sachverhalte." (Wild 2013, 78)

When the bare teleosemantic theory has been spent, the central task for a theory of intentional representation has not yet begun. Teleosemantic theories are piggyback theories. They must ride on more basic theories of representation, perhaps causal theories, or picture theories, or informational theories, or some combination of these.“ (Millikan 2004, 66 zitiert nach Shea 2013, 72f)¹²⁶

Hinsichtlich dieser Interpretation möchte ich mich schlicht der Auffassung von Peter Godfrey-Smith und Nicolas Shea anschließen. Beide sind der Ansicht, dass das Bestehen der Isomorphierelation keine explanatorische Rolle bei der Gehaltsbestimmung nach Millikan spielt oder zumindest, dass dies eine unvorteilhafte Interpretation für Millikan ist.

“So in my view, Millikan's relation of ‚mapping‘ is post hoc, as far as explanations of success are concerned. It is not a resource specifiable in advance, which can play an explanatory role.” (Godfrey-Smith 1996, 186)

“Her view is not, as earlier views were, an attempt to explain representation with the aid of an antecedently existing relation of picturing, isomorphism or mapping between signs and the world.” (Godfrey-Smith 1996, 187)

„However, a better interpretation is that Millikan is not taking the existence of a preexisting isomorphism to be a substantive constraint on the theory of content.“ (Shea 2013, 73)

Das Problem mit der reinen Isomorphie-Relation als Basis für die Bestimmung des Gehalts ist, dass sie viel zu beliebig ist (Vgl. Godfrey-Smith 1996, 184-187). Dies hatten wir auch im Abschnitt zur mapping function festgestellt. Damit irgendeine Isomorphie-Relation besteht, reicht es aus, dass die Anzahl der Elemente aus zwei Bereichen gleich ist. Die Behauptung, dass zwischen zwei Bereichen ein Isomorphismus besteht, ist also extrem schwach. So besteht, wie gesagt, zwischen den acht Stiften in meiner Tasche und den Planeten im Sonnensystem bei geeigneter Zuordnung eine Isomorphie. Oder wie Goodman es ausdrückt hat: „Regularities are where you find them, and you can find them anywhere.“ (Goodman 1983 [1954], 82).

Dementsprechend betonen Peter Godfrey-Smith und Nicolas Shea, dass die relevante Isomorphie-Relation nachträglich spezifiziert wird. Sie ist das Resultat eines Selektionsprozesses (Vgl. hierzu auch Abschnitt 2.1.1.3.8) und Teil der Normalen Erklärungen der Eigenfunktionen von Producer, Icon und Consumer.

„It would get things precisely the wrong way round to claim that there being an isomorphism is a substantive constraint on the theory of content. Instead, Millikan's account of content delivers an isomorphism. It may be relatively benign to talk as if it is the isomorphism as such, rather than a corresponding natural relation, which explains successful behaviour. But that leads to the second, more serious mistake, which is to claim that the existence of isomorphisms is part of what makes it the case that representations represent as they do — that they are prior to the theory of content.“ (Shea 2013, 72)

„[...] it would be a mistake to think that some isomorphism between representations and represented is part of an explanation of why representations have the content they do. Rather, a Normal explanation will advert to physical properties of the system of producer and consumer which make it the case that representations have the content they do — the theory picks out a special isomorphism, rather than being based on the existence of an isomorphism.“ (Shea 2013, 76)

Millikan stimmt Shea, der dieser Problematik in *Millikan and Her Critics* einen ganzen Aufsatz widmet, letztlich zu.

¹²⁶ Shea führt noch eine weitere Stelle an Millikan (2003, 1).

„[...] Shea is right [...] that there is teleology in the analyses prior to the account of error.“ (Millikan 2013c, 82)

Nun gilt es im nächsten Abschnitt zu untersuchen, wie Millikan im Rahmen ihrer Theorie die Möglichkeit zur Missrepräsentation erklärt.

2.2 Missrepräsentation bei Millikan

2.2.1 Zwei Lösungsansätze

Im folgenden Teil der Arbeit wird es darum gehen, genau nachzuvollziehen, wie Millikan das Auftreten von Missrepräsentationen in ihrer Theorie erklärt. Vorerst soll in diesem Abschnitt aber nur dargelegt werden, dass Millikans Biosemantik genaugenommen zwei separate Ansätze enthält, um das Problem der Missrepräsentation in den Griff zu bekommen. Der eine Ansatz beruht *auf der Möglichkeit missgebildeter Funktionsträger und den daraus resultierenden Fehlfunktionen* und der andere Ansatz beruht *auf abNormalen Umweltbedingungen*. Da dies sonst kaum beachtet wird, ist es ein wichtiger Beitrag dieser Arbeit, diese beiden Ansätze klar auseinanderzuhalten.

Was in dieser Arbeit unter einer Missrepräsentation verstanden werden soll, hatten wir in Abschnitt 1.4.3. festgehalten. Dort hatten wir gesagt, dass wir die folgende Situation als paradigmatischen Fall von Missrepräsentation betrachten wollen:

Es ist kein Fressfeind anwesend, dennoch wird ein Informationsträger produziert, der Anwesenheit von Fressfeinden repräsentiert, und der Organismus ergreift die Flucht.

Nachdem wir nun Millikans Theorie eingehend betrachtet haben, können wir in Bezug auf ihre Theorie genauer sagen, dass ein Icon eine Missrepräsentation ist, wenn diejenige Umweltbedingung (Normale Bedingung), die vorliegen müsste, damit der Consumer seine (adaptierte) Eigenfunktion erfüllen kann, wenn er das Icon als unmittelbaren Adaptor nimmt (= Gehalt des Icons), nicht tatsächlich vorliegt (Vgl. Abschnitt 2.1.2).

Die Frage, die uns für den Rest der Arbeit beschäftigen wird, ist nun, ob Millikans Theorie die Ressourcen hat das Auftreten solcher Fälle theoretisch zu erfassen. Die Grundidee teleosemantischer Ansätze, zu denen auch Millikans Biosemantik zählt, ist es, wie bereits erläutert (Vgl. Abschnitt 1.3.4), das Auftreten von Missrepräsentationen durch das Auftreten von Fehlfunktionen zu erklären.

Wenn wir von unserem alltäglichen Verständnis von Funktionen ausgehen, dann ist das Auftreten von Fehlfunktionen eine Banalität, die keiner Erklärung bedarf. Weniger klar ist hingegen, ob dies auch für Millikans Begriff der Eigenfunktion unter Wahrung der naturalistischen Ansprüche gilt.

Bei Millikan finden sich, wie oben bereits erwähnt, eigentlich zwei Ansätze um Fehlfunktionen theoretisch zu erfassen:

- a) Fehlgehen der Funktion aufgrund von abNormalen Umweltbedingungen
- b) Fehlgehen der Funktion aufgrund von missgebildeten Funktionsträgern

Diese Doppelstrategie bleibt aber sowohl bei Millikan als auch in der Sekundärliteratur eher implizit. Dies zeigt sich auch daran, dass Millikan manchmal auf beide Alternativen verweist (missgebildete Funktionsträger *oder* auf abNormale Umweltbedingungen: 1984, 29, 34; 1986, 53, 69f.; 1989a, 294f; 1993a, 5; 2002, 117; 2008, 99), aber eben häufig auch nur eine der beiden Alternativen erwähnt (missgebildete Funktionsträger: 1984, 17; 1989a, 294; abNormale Umweltbedingungen: 1986, 59; 1993b, 130; 1993c, 176). Ähnliches gilt für die Millikanrezeption (beide Alternativen: Detel 2001, 614; Wild 2010b, 890, missgebildete

Funktionsträger: Shea 2005, 13; Wild 2010a, 163, 212-216; abNormale Umweltbedingungen: Samson und Detel 2002, 113)¹²⁷.

Entsprechend der teleosemantischen Idee ergeben sich aus den zwei Ansätzen für das Zustandekommen von Fehlfunktionen auch zwei Ansätze zur Erklärung von Missrepräsentationen:

- a) Missrepräsentationen aufgrund von abNormalen Umweltbedingungen
- b) Missrepräsentationen aufgrund von missgebildeten Funktionsträgern

In den folgenden Abschnitten werde ich die beiden Ansätze separat behandeln und dabei wie folgt vorgehen:

1. Missrepräsentationen aufgrund von abNormalen Umweltbedingungen
 - a. Erläuterung, wie das Fehlgehen von Funktion aufgrund von abNormalen Umweltbedingungen zustande kommt
 - b. Erläuterung, wie Missrepräsentation aufgrund von abNormalen Umweltbedingungen zustande kommt
2. Missrepräsentationen aufgrund von missgebildeten Funktionsträgern
 - a. Erläuterung, wie das Fehlgehen von Funktion aufgrund von missgebildeten Funktionsträgern zustande kommt
 - b. Erläuterung, wie Missrepräsentation aufgrund von missgebildeten Funktionsträgern zustande kommt

¹²⁷ Bei anderen teleosemantischen Theorien herrscht eine ähnliche Vielfalt (beides: Godfrey-Smith 1989, 541; Neander 2012; Neander (in Vorbereitung), Kapitel 4, 3; missgebildete Funktionsträger: Dretske 1986, 22; Neander 1991, 180; Neander 1995, 117, 131; Neander 2006, 184; abNormale Umweltbedingungen: Dretske 1986, 27-32).

2.2.2 AbNormale Bedingungen

2.2.2.1 Umweltbedingte Fehlfunktion

In diesem Abschnitt geht es darum zu erläutern, was nach Millikan ein Verfehlen der Funktion aufgrund von abNormalen Bedingungen sein soll.

Wir können uns dies am besten an dem uns bereits bekannten Beispiel des Industriemelanismus aus Abschnitt 2.1.1.3.4 klar machen. Dort hatten wir gesagt, dass zu den Normalen Bedingungen für die Ausübung der Tarnfunktion der dunklen Außenhülle einer Motte mindestens zwei Bedingungen gehören: a) die Motte befindet sich auf dunklem Untergrund und b) ein Fressfeind ist anwesend. Dementsprechend liegen hier definitiv abNormale Bedingungen vor, wenn entweder der Untergrund der Motte hell ist oder kein Fressfeind anwesend ist oder beides der Fall ist.

Angenommen wir beobachten nun wie eine dunkle Motte, die auf einem hellen Untergrund sitzt (abNormale Bedingung), ins Blickfeld eines Fressfeindes gerät (Normale Bedingung) und kurz darauf gefressen wird. In diesem Fall scheint die dunkle Außenhülle der Motte ihre Tarnfunktion aufgrund der abNormalen Bedingung (heller Untergrund) zu verfehlen. Weniger plausibel ist hingegen zu behaupten, dass die dunkle Außenhülle der Motte ihre Tarnfunktion verfehlt, wenn wir beobachten, dass die Motte zwar auf dunklen Untergrund sitzt (Normale Bedingung), aber sich nicht im Blickfeld eines Fressfeindes befindet (abNormale Bedingung). In diesem Fall wird zwar kein Tarneffekt bewirkt, denn für diesen ist die Anwesenheit eines Fressfeindes nötig¹²⁸, aber dennoch sind wir wahrscheinlich intuitiv nicht geneigt dies als Verfehlen der Eigenfunktion der dunklen Außenhülle zu betrachten. Wir würden wahrscheinlich viel eher behaupten, dass es sich gar nicht um einen „Anwendungsfall“ der Eigenfunktion handelt, d.h. es handelt sich um eine Situation, in der die Eigenfunktion weder erfüllt noch verfehlt werden kann.

Diese Intuition stellt zunächst kein Problem für Millikan dar, denn sie behauptet nicht, dass abNormale Bedingungen generell zum Verfehlen von Funktionen führen, sondern nur, dass ein Verfehlen der Funktion aufgrund von abNormalen Bedingungen möglich ist. Trotzdem wüsste man gern genauer, wann abNormale Bedingungen für ein Verfehlen verantwortlich sind und wann nicht. Gibt es eventuell eine bestimmte Sorte von abNormalen Bedingungen, die ein Verfehlen von Funktionen hervorrufen?

Um diese Frage zu beantworten, mag es hilfreich sein sich ein weiteres Beispiel anzuschauen. Das Beispiel stammt von Millikan selbst und betrifft die Funktion der Schweißdrüse durch Absonderung von Schweiß, die Körpertemperatur zu senken. Millikan nennt zwei Normale Bedingungen für das Ausüben der Funktion der Schweißdrüsen, erstens muss eine erhöhte Körpertemperatur vorliegen und zweitens muss die Luftfeuchtigkeit unter 100 % sein.

„Now a Normal condition for sweating to cause cooling of the body is that the humidity is less than 100%. [...] If we say, using slightly more precision, that a proper function of the sweating is to help return the body to its Normal temperature, then we see that the body's being overheated is another Normal condition for the sweating's serving its function.“ (Millikan 1986, 69)

Es gibt also auch hier wieder zwei theoretische Möglichkeiten für ein Verfehlen der Funktion aufgrund abNormaler Bedingungen. Entweder die Luftfeuchtigkeit ist unter 100% oder die

¹²⁸ Zumindest wenn der Tarneffekt als ein Nicht-gesehen-werden-durch-den-Fressfeind und nicht als ein Nicht-gesehen-werden-können verstanden wird.

Körpertemperatur ist nicht erhöht¹²⁹. Im ersten Fall sondert die Schweißdrüse zwar Schweiß ab, kann aber aufgrund der hohen Luftfeuchtigkeit (abNormale Bedingung) kein Absenken der erhöhten Körpertemperatur (Normale Bedingung) bewirken. Die Schweißdrüse scheint in diesem Fall rein intuitiv ihre Funktion zu verfehlen. Im zweiten Fall sondert die Schweißdrüse bei einer Luftfeuchtigkeit von unter 100% (Normale Bedingung) keinen Schweiß ab und senkt somit auch nicht die ohnehin nicht erhöhte Körpertemperatur (abNormale Bedingung). Das Verhalten der Schweißdrüse ist dabei schwerlich als ein Verfehlen ihrer Funktion zu interpretieren. Vielmehr würden wir, wie schon beim Motten-Beispiel, sagen, dass hier kein „Anwendungsfall“ für die Ausübung der Funktion vorliegt.

Betrachten wir noch ein weiteres Beispiel, bevor wir versuchen unsere Überlegungen zusammenzufassen. Das von Fred Dretske (1986) stammende Beispiel der Magnetotaxis einiger Bakterien haben wir bereits in Abschnitt 1.3.4.2 kennengelernt. Zur Erinnerung:

„Some marine bacteria have internal magnets (called magnetosomes) that function like compass needles, aligning themselves (and, as a result, the bacteria) parallel to earth magnetic field. Since these magnetic lines incline downwards (towards geomagnetic north) in the northern hemisphere (upwards in the southern hemisphere), bacteria in the northern hemisphere, oriented by their internal magnetosomes, propel themselves towards geomagnetic north. The survival value of magnetotaxis (as this sensory mechanism is called) is not obvious, but it is reasonable to suppose that it functions so as to enable the bacteria to avoid surface water. Since the organisms are capable of living only in the absence of oxygen, movements towards geomagnetic north will take the bacteria away from oxygen-rich surface water and towards the comparatively oxygen-free sediment at the bottom.“ (Dretske 1986, 26)

Zwei Bedingungen scheinen für die Erklärung der Funktionsweise der Magnetosomen besonders wichtig zu sein. Die erste Bedingung ist, dass sauerstoffreiches Wasser für die Bakterien giftig ist *und* dass es sauerstoffreiches Wasser in der Umgebung des Bakteriums gibt (Oberflächenwasser ist sauerstoffreich, während Wasser am Grund sauerstofffrei ist). Die zweite Bedingung ist, dass ein Schwimmen zum geomagnetischen Norden zugleich ein Schwimmen in Richtung Gewässergrund ist (das Bakterium muss sich also in der nördlichen Hemisphäre befinden und es darf kein Magnet seiner Nähe sein). Beide Bedingungen sind also Normale Bedingungen dafür, dass die Magnetosomen ihre Funktion ausüben können. Daraus ergeben sich wiederum zwei Möglichkeiten für das Verfehlen der Funktion aufgrund von abNormalen Bedingungen. Bei der ersten Möglichkeit bewegt sich ein Bakterium in einer Umgebung mit für ihn giftigem, sauerstoffreichem Oberflächenwasser (Normale Bedingung) und es befindet sich in der südlichen Hemisphäre (abNormale Bedingung), was dazu führt, dass es in den Tod schwimmt. In diesem Fall würden wir wahrscheinlich intuitiv sagen, dass die Magnetosomen ihre Funktion aufgrund der abNormalen Umweltbedingung (südliche Hemisphäre) verfehlt haben¹³⁰. Bei der zweiten Möglichkeit bewegt sich ein Bakterium in einer Umgebung, in der es überhaupt kein sauerstoffreiches Wasser gibt (abNormale Bedingung)¹³¹, und es befindet sich in der nördlichen Hemisphäre (Normale Bedingung), was dazu führt, dass es zum Gewässergrund schwimmt. In diesem Fall wird kaum jemand davon sprechen wollen, dass die Magnetosomen ihre Funktion verfehlen. Viel eher würden wir

¹²⁹ Die dritte Möglichkeit, dass beides der Fall ist, wollen wir wie beim vorangegangenen Beispiel nicht näher betrachten.

¹³⁰ Wir hatten aber schon in Abschnitt 1.3.4.2 mit Dretske darauf hingewiesen, dass diese Einschätzung unmittelbar davon abhängt wie wir die Funktion der Magnetosomen beschreiben. Wenn die Funktion der Magnetosomen lediglich darin besteht die Schwimmrichtung des Bakteriums nach dem vorherrschenden Magnetfeld auszurichten, dann verfehlen die Magnetosomen ihre Funktion im gerade geschilderten Fall nicht, sondern erfüllen sie (Vgl. Dretske 1986, 29-32). Auf dieses Problem werden wir noch zurückkommen.

¹³¹ Oder für das Bakterium ist sauerstoff nicht giftig.

sagen, dass sie ihre Funktion erfüllen oder dass es sich nicht um einen „Anwendungsfall“ für die Funktion handelt.

Wir wollen nun versuchen unsere Überlegungen zu den einzelnen Beispielen zusammenfassend zu verallgemeinern. Wir hatten festgestellt, dass es in allen drei Beispielen Bedingungen (Fressfeind anwesend, überhöhte Körpertemperatur, Sauerstoff in der Umgebung) gab, auf die der jeweilige Organismus (Motte, Mensch, Bakterium) reagieren musste, weil er sonst nicht überlebt. Es drohte: Gefressenwerden, Hitzetod, tödliche Vergiftung. Alle Organismen haben nun einen Funktionsträger (Außenhülle, Schweißdrüse, Magnetosom) dessen Aufgabe es ist, das jeweils drohende Resultat zu verhindern. Der Funktionsträger erfüllt seine Funktion also, wenn er einen bestimmten Effekt (Nichtgefressenwerden, Absenken der Körpertemperatur, Vermeidung von sauerstoffreichem Wasser) bewirkt. Diesen Effekt wiederum bewirkt der Funktionsträger auf eine bestimmte Weise (durch Tarnung, durch Schweißabsonderung, durch Schwimmen zum geomagnetischen Norden). Damit der Effekt wiederum auf die bestimmte Weise erreicht werden kann, müssen weitere Bedingungen (Motte sitzt auf dunklem Untergrund, die Luftfeuchtigkeit liegt unter 100%, das Bakterium befindet sich in der nördlichen Hemisphäre) erfüllt sein. D.h. wir können unter den Normalen Bedingungen zwei Gruppen unterscheiden:

Zur *ersten Gruppe* gehören diejenigen Bedingungen, aufgrund derer überhaupt ein „Handlungsbedarf“ seitens des Organismus besteht.

Zur *zweiten Gruppe* gehören diejenigen Bedingungen, die für die spezifische Problemlösung des Organismus' gegeben sein müssen.

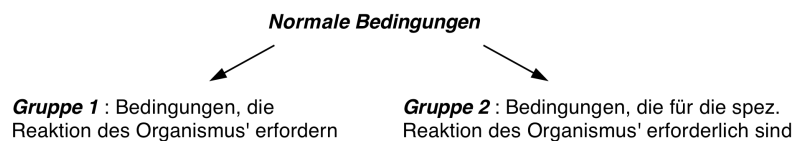


Abbildung 48: Zwei Gruppen von Normalen Bedingungen

Insgesamt können wir diese Ergebnisse in der folgenden Tabelle zusammenfassen.

Beispiel	Normale Bedingungen		Effekte	
	Gruppe 1	Gruppe 2	(Mittel)	(Ziel)
Motte (dunkel)	Freßfeind anwesend	Motte auf dunklem Untergrund	kein Kontrast zur Umgebung (Tarnung)	nicht gefressen werden
Schweißdrüse	erhöhte Körpertemperatur	Luftfeuchtigkeit unter 100%	Schweiß absondern	Körpertemperatur absenken
Bakterium (Magnetotaxis)	O ₂ -reiches Oberflächenwasser	Bakterium auf nördl. Hemisphäre	Schwimmen zum magnetischen Norden	Vermeiden von O ₂ -reichem Wasser

Abbildung 49: Unterscheidung der zwei Gruppen von Normalen Bedingungen und verschiedener Effekte (Mittel/Ziel) anhand der drei behandelten Beispiele

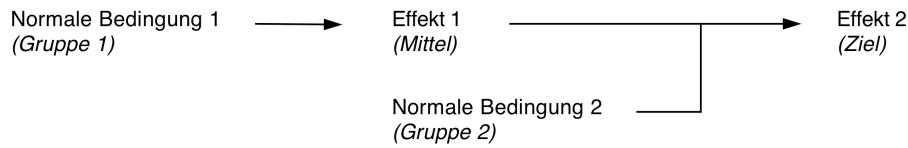


Abbildung 50: Wirkungsbeziehung zwischen den zwei Gruppen von Normalen Bedingungen und den Effekten (Mittel/Ziel)

Die Beispiele legen nahe, dass nur ein Abweichen von der zweiten Gruppe von abNormalen Bedingungen überhaupt den Eindruck vermitteln kann, dass die jeweiligen Funktionsträger in diesen Situationen ihre Funktion verfehlen. Wir wollen auch Millikan so verstehen, dass sie sich bei einem Verfehlen aufgrund von abNormalen Umweltbedingungen immer auf ein Abweichen von der zweiten Gruppe von Normalen Bedingungen bezieht.

So in etwa könnte Millikans Erklärung für ein Verfehlen der Funktion aufgrund von abNormalen Umweltbedingungen ausformuliert lauten:

Das Verfehlen der Funktion entsteht dadurch, dass der Funktionsträger, dessen Aufgabe es ist ein Resultat, das die Folge einer bestimmten Bedrohung/Situation ist, zu verhindern, nicht in der Lage ist das Resultat bei Vorliegen der Bedrohung/Situation tatsächlich zu verhindern. Es misslingt ihm, nicht weil er defekt ist, sondern weil andere Bedingungen, die für seine spezifische Arbeitsweise notwendig sind und sich seiner Kontrolle entziehen, nicht erfüllt sind.

Dementsprechend gilt für Millikan:

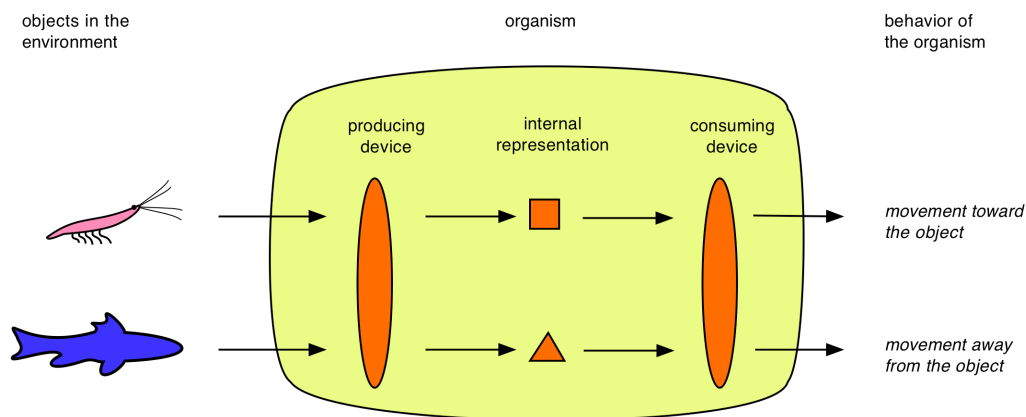
- (1) Die dunkle Außenhülle einer Motte verfehlt ihre Funktion, wenn die Motte auf hellem Untergrund sitzend von einem Fressfeind entdeckt und gefressen wird.
- (2) Eine Schweißdrüse verfehlt ihre Funktion, wenn sie bei 100% Luftfeuchtigkeit keine Absenkung der erhöhten Körpertemperatur durch Schweißabsonderung bewirkt.
- (3) Die Magnetosome eines Bakteriums verfehlen ihre Funktion, wenn sie auf der südlichen Hemisphäre das Bakterium nicht vom sauerstoffreichen Oberflächenwasser fernhalten, indem sie es Richtung geomagnetischer Norden schwimmen lassen.

Wir haben nun hinreichend genau nachvollzogen, wie Millikan das Verfehlen der Funktion durch einen Funktionsträger auf abNormale Umweltbedingungen zurückführt, und damit das Ziel dieses Abschnittes erreicht.

2.2.2.2 Umweltbedingte Missrepräsentation

In diesem Abschnitt geht es darum nachzuvollziehen, wie nach Millikan Missrepräsentationen aufgrund von abNormalen Umweltbedingungen zustande kommen. Da der teleosemantischen Grundidee nach dem Auftreten von Missrepräsentationen das Verfehlen einer Funktion zugrunde liegt, geht es genauer gesagt darum, zu verstehen, wie das Verfehlen einer Funktion aufgrund von abNormalen Bedingungen zu Missrepräsentationen führen kann.

Betrachten wir zunächst wieder unser fiktives Wasserlebewesen. Für das Zustandekommen von Missrepräsentationen ist vor allem der producing device relevant, denn er produziert die Repräsentationen.



Was wären nun die Normalen Bedingungen für die Eigenfunktion des producing device? Wir können hier wieder wie im vorangegangenen Abschnitt zwischen zwei Gruppen von Normalen Bedingungen unterscheiden:

Gruppe 1: Bedingungen, aufgrund derer überhaupt ein Handlungsbedarf seitens des Organismus besteht.

Gruppe 2: Bedingungen, die für die spezifische Problemlösung des Organismus' gegeben sein müssen.

Zur ersten Gruppe gehört im Fall des producing devices ganz sicher die Anwesenheit entweder von Nahrung oder von Fressfeinden. Was zur zweiten Gruppe gehören würde, hängt von der Arbeitsweise des producing devices ab. Nehmen wir an, der producing device ist ein lichtsensitives Organ, dann würde zu den Normalen Bedingungen seiner Funktion sicherlich eine gewisse Lichtintensität gehören, die nicht unterschritten werden darf. Im vorangegangenen Abschnitt hatten wir zudem festgestellt, dass der Eindruck des Verfehlens der Funktion vor allem dann entsteht, wenn die Normalen Bedingungen der Gruppe 1 erfüllt sind aber die Normalen Bedingungen der Gruppe 2 nicht.

Eine solche Situation würde etwa vorliegen, wenn ein Fressfeind anwesend ist (Normale Bedingung der Gruppe 1 erfüllt) aber eine zu geringe Lichtintensität herrscht (Normale Bedingung der Gruppe 2 nicht erfüllt). In dieser Situation könnte man analog zum vorangegangenen Abschnitt vermuten, dass der producing device seine Funktion verfehlen wird. Daraus folgt aber nicht, dass er auch eine Missrepräsentation hervorbringt. Im Gegenteil, es ist sogar wahrscheinlicher, dass der producing device in diesem Fall überhaupt nicht reagieren würde (siehe Abbildung 51).

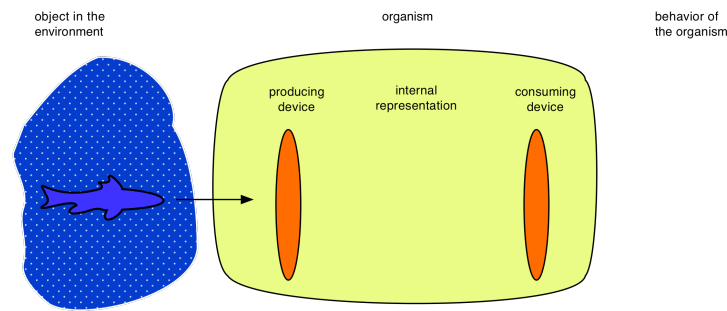


Abbildung 51: Producing device produziert aufgrund Abnormaler Bedingungen (Gruppe 2) gar keine interne Repräsentation

Es ist es aber durchaus nicht undenkbar, dass der producing device unter ganz spezifischen Lichtverhältnissen bei Anwesenheit eines Fressfeindes systematisch das Vierecks-Icon und damit eine Missrepräsentation produziert (siehe Abbildung 52).

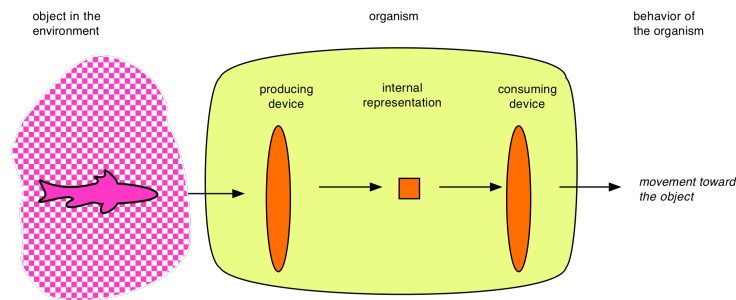


Abbildung 52: Producing device produziert aufgrund Abnormaler Bedingungen (Gruppe 2) ein Vierecks-Icon anstatt eines Dreiecks-Icons

Insgesamt ließe sich so jedenfalls einigermaßen plausibel das Auftreten von Missrepräsentationen unter abNormalen Umweltbedingungen bei Millikan rekonstruieren.

Tatsächlich spielt aber der durch Abbildung 52 dargestellte Fall in der Diskussion um das Problem der Missrepräsentation eher eine untergeordnete Rolle¹³². Gerade im Zusammenhang mit dem Disjunktionsproblem (Vgl. Abschnitt 1.3.2) hat man sich eher am Fall des Frosches, der nach vorbei fliegenden Plastik-Pellets schnappt, abgearbeitet¹³³ – einen Fall also, in dem ein Organismus durch eine Attrappe „getäuscht“ wird. In Bezug auf unser fiktives Wasserlebewesen können wir uns einen solchen Fall vorstellen, wenn wir z.B. annehmen, dass Fisch-Attrappen in der Umgebung des Wasserlebewesens ausgesetzt wurden sind. Die Begegnung mit einer solchen Attrappe könnte dann dazu führen, dass der producing device ein Dreiecks-Icon produziert, auf das der consuming device mit dem Einleiten einer Fluchtbewegung reagiert (siehe Abbildung 53).

¹³² Nur Fodor verweist gelegentlich implizit auf die Lichtverhältnisse, wenn z.B. von “being (the right sort of) cat-on-a-dark-night” (Fodor 1990, 59) die Rede ist.

¹³³ Vgl. Sterelny (1990, 124ff), Millikan (1991), Price (2001, 80), Neander (1995, 130), Jakob (1997, 134), Dennett (1987, 320), Rowlands (2006, 130ff), Wild (2010, 371), Agar (1993), Fodor (1990), Schulte (2012). “This story was originally inspired by Lettvin et al.’s (1959) experiments on leopard frogs (*Rana pipiens*), but has since developed a life of its own” (Schulte 2012, 484).

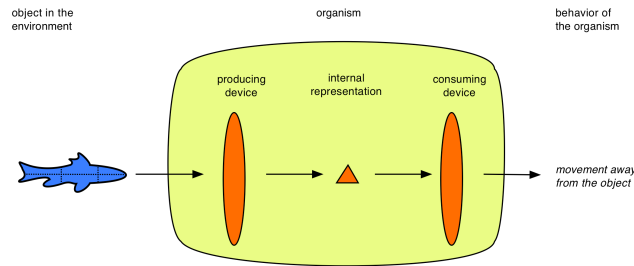


Abbildung 53: Producing device produziert aufgrund Abnormaler Bedingungen (Gruppe 1) eine Dreiecks-Icon anstatt gar keine Repräsentation hervorzubringen

Das Dreiecks-Icon könnte man dann als Missrepräsentation verstehen, die unter abNormalen Bedingungen (Abwesenheit von Fressfeind¹³⁴) zustande kommt.

Interessanterweise entspricht der durch Abbildung 53 dargestellte Fall, aber eher denjenigen Fällen, bei denen wir im vorgegangen Abschnitt intuitiv gesagt hatten, es handele sich hier um Situationen, die gar keinen Anwendungsfall der entsprechenden Funktion darstellen. Diese Intuition hatten wir immer dann, wenn abNormale Bedingungen der Gruppe 1 vorlagen und genau das ist auch hier der Fall.

Später werden wir noch untersuchen, ob abNormale Bedingungen nicht generell dazu führen, dass die Situation kein Anwendungsfall der Funktion ist. Sollte dies zutreffen, dann wäre Millikans Lösung des Problems der Missrepräsentation über abNormale Bedingungen sehr fragwürdig (Vgl. Abschnitt 3.2.9).

¹³⁴ und Nahrung.

2.2.3 Missgebildete Funktionsträger

2.2.3.1 Missbildungsbedingte Fehlfunktion

In diesem Abschnitt geht es darum nachzuvollziehen, wie nach Millikan ein Verfehlen der Funktion aufgrund von missgebildeten Funktionsträgern zustande kommt. Die eigentliche Schwierigkeit hierbei besteht darin, theoretisch zu begründen, wieso etwas trotz abweichender physischer Merkmale (die Missbildung) als zu einer bestimmten Gruppe von Funktionsträgern zugehörig verstanden werden soll.

Schauen wir uns zunächst die Stelle an, in der Millikan behauptet, dass es für das Haben einer Eigenfunktion nicht einmal notwendig sei, dass der Funktionsträger diese überhaupt ausüben *könne*, eben weil es missgebildete Mitglieder reproduktiv etablierter Familien gäbe:

„It is not necessary that a device even be capable of serving any direct proper function of it. Some members of reproductively established families are malformed and unable for this reason to perform properly. Having a proper function depends upon the *history* of the device that has it, not upon its form of dispositions.“ (Millikan 1984, 29)

D.h. etwas kann eine Funktion haben und unter gar keinen Bedingungen in der Lage sein diese auszuüben. Dies wäre mit einer dispositionalen Funktionsauffassung, die uns hier als Kontrastfolie dienen mag, nicht vereinbar, denn unter allen möglichen Bedingungen sind natürlich auch die für die Disposition relevanten und unter denen muss sich die Disposition manifestieren (die Funktion ausgeübt werden) oder das Objekt hat die Disposition (Funktion) nicht. Nach Millikan trifft aber genau dies auf missgebildete REF-Mitglieder zu, sie haben eine Eigenfunktion können diese aber auch unter den für die Eigenfunktion relevanten Bedingungen, den Normalen Bedingungen, nicht ausüben. Mit anderen Worten: Bei missgebildeten REF-Mitgliedern ist ein Verfehlen der Funktion sogar unter Normalen Bedingungen möglich, daher ist die Begründung der Möglichkeit von missgebildeten REF-Mitgliedern für Millikan von größter Bedeutung.

Millikans Aussagen zu missgebildeten REF-Mitgliedern scheinen sich zunächst *nur* auf HOREF-Mitglieder zu beziehen. Das legen ihre Beispiele, Auge und Herz, nahe (Vgl. Millikan 1984, 25, 29). Zudem fügt sie bei der Definition der reproduktiv etablierten Familie höherer Ordnung (Vgl. Abschnitt 2.1.1.3.5) eine dritte Bedingung hinzu unter der Objekte ebenfalls als Mitglieder einer HOREF gelten sollen, wobei die Bedingung explizit Raum für die Möglichkeit missgebildeter REF-Mitglieder schaffen soll.

„This condition is not an additional condition, but a loosening of conditions (1) und (2) above. The point of this new condition is to make room for malformed members of reproductively established families.

- (3) If anything x (a) has been produced by a device a direct proper function of which is to produce a member or members of a higher-order reproductively established family R , and (b) is in some respects like Normal members of R because (c) it has been produced in accordance with an explanation that approximates in some (undefined) degree to a Normal explanation for production of members of R , then x is a member of R .“ (Millikan 1984, 25)

Versuchen wir uns Millikans Idee an einem abstrakten Schema verständlicher zu machen. Vorausgesetzt die Mitglieder der HOREF A haben die Eigenfunktion Mitglieder der HOREF B

zu produzieren und Mitglieder der HOREF B haben die Eigenfunktion einen Effekt C zu produzieren, dann wäre Folgendes möglich:

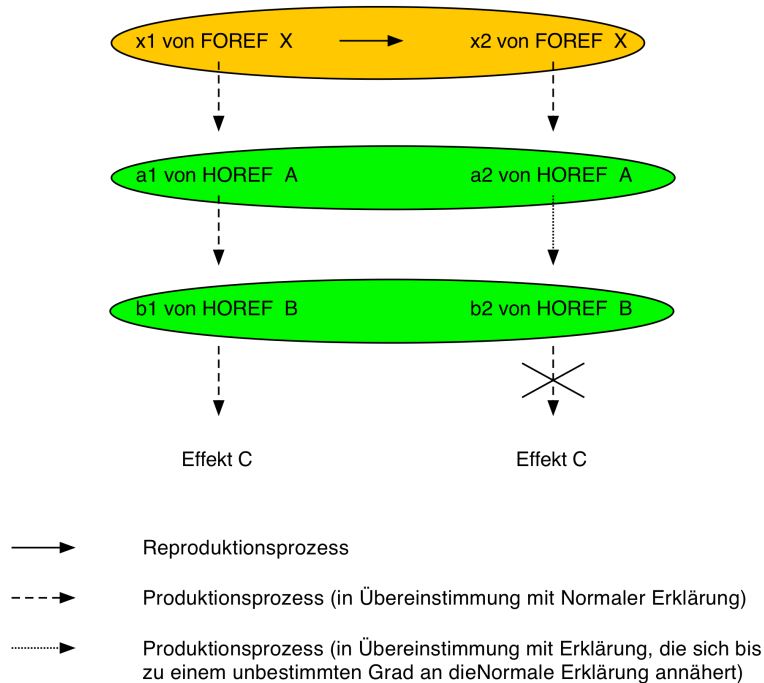


Abbildung 54: Schematische Darstellung des Verfehlers einer Funktion aufgrund eines missgebildeten HOREF-Mitgliedes

Das Objekt b_2 ist ein Mitglied der HOREF B, weil es von einem REF A-Mitglied, nämlich a_2 produziert wurde und es die Aufgabe von HOREF A-Mitgliedern ist HOREF B-Mitglieder zu produzieren¹³⁵. HOREF B-Mitglieder haben die Funktion den Effekt C zu produzieren, also hat b_2 die Funktion den Effekt C zu produzieren, obwohl es dazu nicht in der Lage ist. Der Gedankengang hat dabei die Form eines Syllogismus:

All B's are C's	All members of HOREF B have the function to produce the effect C
b is a B	b_2 is a members of HOREF B
<hr/>	<hr/>
b is a C	b_2 has the function to produce the effect C

¹³⁵ Das ist nur die Bedingung (a) aus dem Millikanzitat. Die anderen beiden Bedingungen, vor allem (c), sorgen dafür, dass nicht jedes beliebige Produkt eines HOREF A-Mitglieds als HOREF B-Mitglied gilt (Vgl. auch Abschnitt 2.1.1.3.4). Millikan macht aber keine genauen Angaben wie REF B-Mitglieder von beliebigen REF A-Produkten unterschieden werden können. Dass die Übergänge dabei fließend sind und es somit Grenzfälle gibt, scheint mir akzeptabel. Und Millikan gibt diese Vagheit auch zu: „The vagueness of the question, in some cases, whether a bit of matter should be called ‚a malformed eye‘ or merely ‚a glob of misplaced organic matter on the forehead‘ is reflected in the vagueness of ‚in some respects like‘ and ‚approximates in some degree to a Normal explanation.‘“ (Millikan 1984, 25) Dass es Grenzfälle gibt, heißt aber nicht, dass man nicht zumindest angeben kann, was ein eindeutiger Fall eines missgebildeten HOREF-Mitgliedes wäre. Aber auch dies macht Millikan nicht.

Millikans Pointe liegt natürlich in der zweiten Prämisse. Die Frage der Mitgliedschaft wird durch die Herkunft entschieden und nicht durch die tatsächlichen Eigenschaften eines Objektes. Aber diese Grundidee wird durch die im vornherein eingebaute Vagheit der oben erwähnten Zusatzbedingung geschwächt. Man kann sich hier fragen, ob die Vagheit nicht zu einer Immunisierung gegen Gegenbeispiele führt und damit die Frage nach der Mitgliedschaft zu einer HOREF zu einer mehr oder weniger subjektiven Angelegenheit wird¹³⁶. Zudem fehlt uns generell oft das exakte Wissen über die Ontogenese von HOREF-Mitglieder. Dieses Wissen ist aber notwendig für die Normale Erklärung in Übereinstimmung mit der ein Objekt produziert werden muss, damit es als ein Mitglied einer bestimmten HOREF gelten kann. Mit anderen Worten: Uns fehlt oft der Standard, dem sich die Erklärung der Produktion eines missgebildeten REF-Mitgliedes bis zu einem unbestimmten Grad annähern soll.

Alles in allen ist also eine gewisse Skepsis gegenüber Millikans Bestimmung von missgebildeten HOREF-Mitgliedern nicht unangebracht. Wir werden später bei der Bewertung von Millikans Lösungen auf diese potentielle Schwachstelle zurückkommen und diese eingehend prüfen (Vgl. Abschnitt 3.2.8).

¹³⁶ Millikan scheint dieses Problem zumindest im Prinzip zuzugeben (siehe Zitat Fn 129).

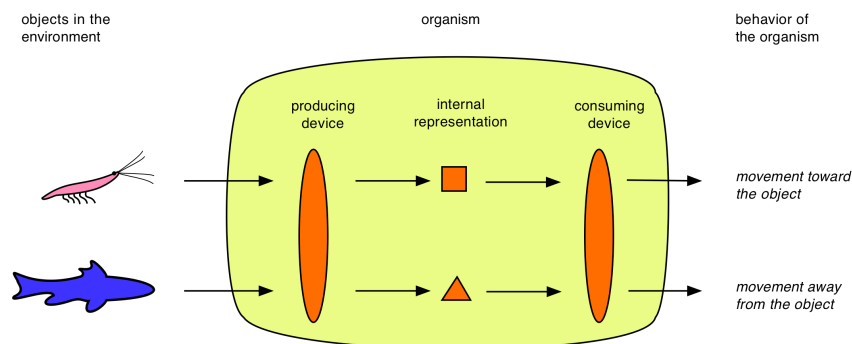
2.2.3.2 Missbildungsbedingte Missrepräsentation

In diesem Abschnitt geht es darum nachzuvollziehen, wie nach Millikan Missrepräsentationen aufgrund von missgebildeten Funktionsträgern zustande kommen.

Explizite und ausführliche Darstellungen dazu sucht man bei Millikan allerdings vergebens. Die beste Darstellung findet man noch in Millikan (1984) und zwar eigenartigerweise an einer Stelle, die sich noch vor Millikans Darlegung ihrer Generellen Zeichentheorie befindet. Dort erläutert Millikan mehr nebenbei, wie ein konkreter Bientanz, der der Bientanz-Syntax entspricht, aufgrund eines defekten (abNormalen) *dance-choreographing mechanisms* missrepräsentiert, weil er die Lage einer Nektarquelle falsch abbildet (Millikan 1984, 42-43).

"A device is maladapted to a particular adaptor if it does not bear the relation it is supposed to bear to that adaptor. An adapted device is also maladapted if it has no real adaptor. [...] Where adapted devices are maladapted, it can happen that one and the same device acquires conflicting proper functions. Suppose that a worker bee spots nectar 2000 feet southeast of the hive but, owing to some abNormality of its dance-choreographing mechanisms or owing to abNormal appearance, say, of the sun, it produces a maladapted dance. The dance it produces displays correct B-mese syntax but, by B-mese rules, is of a kind that should be produced when nectar is north-northwest. This dance has as a derived adapted proper function to move the bee family southeast. But as a member of the reproductively established family of syntactically correct bee dances it has an adapted proper function (adapted to its own concrete form) to move the bee family north-northwest. Also, the interpreter mechanism of any bee that watches the dance acquires the adapted proper function of moving that bee north-northwest, qua being supposed to produce a line of flight that bears such and such a relation to the dance watched. But a further proper function of these devices is to move the bee toward nectar and the nectar is not north-northwest but southeast. Notice that what we would intuitively take to be what is represented in the case of the maladapted bee dance hangs upon the dance's direct proper function, this according with the adapted proper function of the interpreter bee's interpreting mechanisms qua confronted with the dance." (Millikan 1984, 43).

Wir können uns Millikans Vorstellung wieder an unserem fiktiven Wasserlebewesen klar machen.



Wie bei Millikans Bientanzbeispiel ist für das Zustandekommen von Missrepräsentationen vor allem der producing device relevant, denn er produziert die Repräsentationen. In Millikans Beispiel ist der producing device der *dance-choreographing device* und in unserem Beispiel das sensorische System des Wasserlebewesens. Nehmen wir nun an, dass das

sensorische System des Wasserlebewesens aufgrund einer Umwelteinwirkung deformiert wurde und zwar derart, dass es sowohl auf Nahrung (kleine, rötliche Objekte) als auch auf Fressfeinde (grosse, blaue Objekte) in der Umwelt mit der Produktion einer „Vierecks-Repräsentation“ reagiert (siehe Abbildung 55).

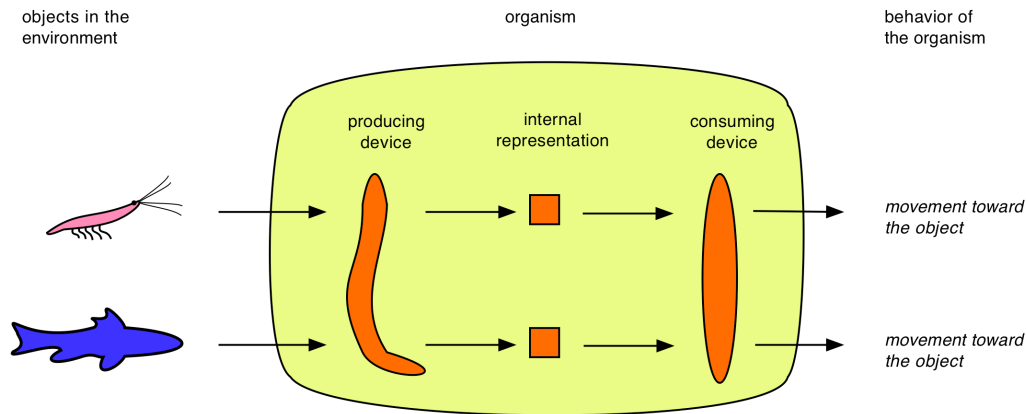


Abbildung 55: Deformierter producing device produziert Vierecks-Icons auch in Gegenwart von Fressfeinden

Der *producing device* dieses Wasserlebewesens produziert also in Gegenwart eines Fressfeindes (grossen, blauen Objekts) eine „Vierecks-Repräsentation“ und somit eine Missrepräsentation, denn eine „Vierecks-Repräsentation“ besagt nach der Abbildungsregel, dass Nahrung (ein kleines, rötliches Objekt) in der Umwelt ist, während es tatsächlich ein Fressfeind (ein grosses, blaues Objekt) ist. Es ist zentral für Millikans Lösungsansatz, wie Missrepräsentation möglich ist, dass auch der defekte *producing device* etwas produziert, das der Syntax des repräsentationalen Systems entspricht und also weiterverarbeitet werden und ein Verhalten erzeugen kann, in unserem Beispiel also „Vierecks-“ oder „Dreiecks-Repräsentationen“. Mit anderen Worten: Ein Verfehlen der Funktion der *producing devices* allein ist nicht hinreichend für das Auftreten einer Missrepräsentation. Dies hatten wir auch schon in Abschnitt 2.2.2.2 festgestellt. Wir können uns dies wiederum an unserem Beispiel verdeutlichen. Angenommen der *producing device* wurde in einer Weise deformiert, in der er eine „Kreis-(Pseudo)Repräsentation“ in Gegenwart eines Fressfeindes (grossen, blauen Objektes) produziert. Die Kreis-Repräsentation ist dann keine Missrepräsentation, sondern eine Pseudo-Repräsentation. Hier läuft zwar auch etwas schief, d.h. der *producing device* verfehlt seine Funktion, aber die „Kreis-(Pseudo)Repräsentation“ bildet keine *andere* Umweltbedingung ab als diejenige, die tatsächlich vorliegt. Da sie nicht der Syntax des repräsentationalen Systems entspricht, kann sie auch nichts abbilden (siehe Abbildung 56).

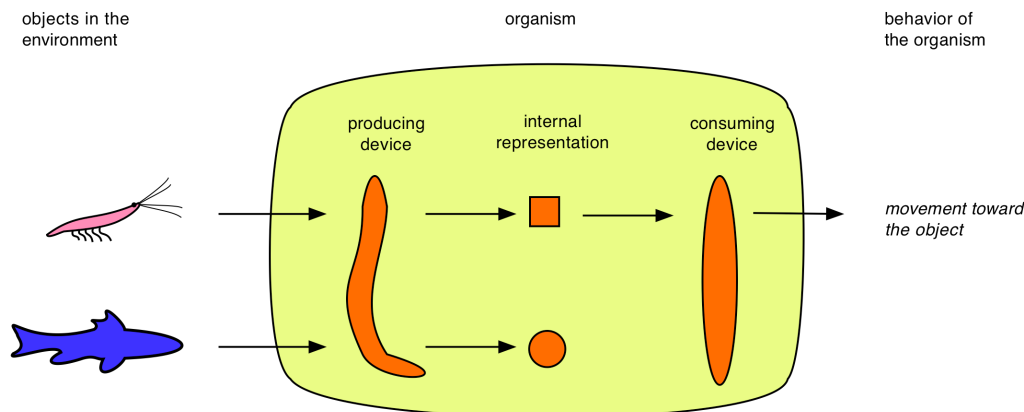


Abbildung 56: Deformierter producing device produziert Kreis-(Pseudo)-Icons in Gegenwart von Fressfeinden

Interessanterweise erlaubt Millikans Erläuterung der Entstehung von Missrepräsentation aufgrund deformierter Funktionsträger prinzipiell eine Lösung des Problems der Missrepräsentation ohne zugleich das Disjunktionsproblem zu lösen. Dies ist wichtig, weil aufgrund der engen Verbindung, die zwischen dem Problem der Missrepräsentation und dem Disjunktionsproblem gesehen wird, der Eindruck entsteht, dass eine Lösung des ersten Problems zwingend eine Lösung des zweiten Problems erfordere. Führen wir uns den Zusammenhang zwischen dem Problem der Missrepräsentation und dem Disjunktionsproblem noch einmal an unserem Beispiel vor Augen (Vgl. auch Abschnitt 1.3.2 und Abschnitt 2.2.2.2): Normalerweise reagiert der Organismus auf Raubfische (angenommen das sind die einzigen grossen blauen Objekte in seiner Umwelt) mit einer „Dreiecks-Repräsentation“ und Flucht. Wir könnten daher die Abbildungsregel mit *Raubfisch – Dreieck* bestimmen. Angenommen ein Verhaltensbiologe baut eine Raubfisch-Attrappe (ebenfalls ein grosses, blaues Objekt) und konfrontiert damit nun unseren Organismus. Der Organismus reagiert auf die Attrappe mit einer „Dreiecks-Repräsentation“ und Flucht (siehe Abbildung 57).

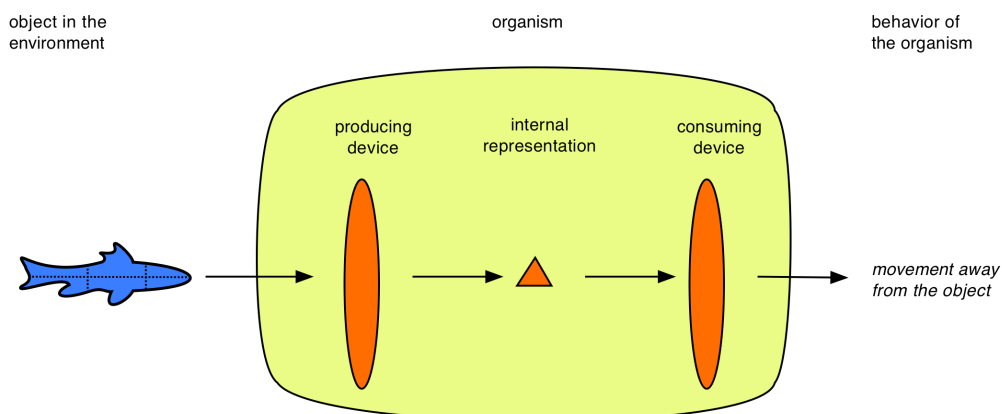


Abbildung 57: Producing device produziert Dreiecks-Icons in Gegenwart einer Fressfeindattrappe

Intuitiv würden wir wahrscheinlich zunächst sagen, der Organismus missrepräsentiere die Attrappe als Raubfisch. Hierbei legen wir aber die Abbildungsregel *Raubfisch – Dreieck*

zugrunde. Wir könnten der „Dreiecks-Repräsentation“ aber genauso gut die Disjunktion „Raubfisch oder Raubfisch-Attrappe“ zuordnen oder „grosses, blaues Objekt“. Wir erhielten dann die Abbildungsregeln *Raubfisch oder Raubfisch-Attrappe – Dreieck* oder *grosses, blaues Objekt – Dreieck* und nach beiden Abbildungsregeln ist die in Abbildung 57 dargestellte Situation kein Fall von Missrepräsentation. Eine Lösung des Disjunktionsproblems besteht dementsprechend darin, eine möglichst gute Begründung für eine der Abbildungsregeln zu finden. Sollte die Begründung zugunsten der ersten Abbildungsregel ausfallen, hat man zugleich das Problem der Missrepräsentation gelöst.

Wenn der Organismus allerdings einen deformierten *producing device* wie in Abbildung 58 hat, dann ist eine solche Lösung unnötig, zumindest wenn es darum geht Missrepräsentationen überhaupt zu ermöglichen. Die „Vierecks-Repräsentation“ und anschließende Annäherung als Reaktion auf die Attrappe ist nach allen vorgeschlagenen Abbildungsregeln eine Missrepräsentation (siehe Abbildung 58).

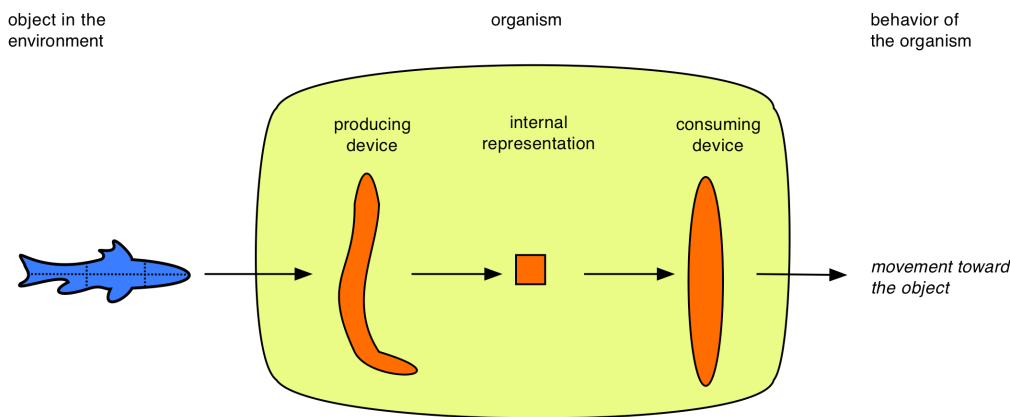


Abbildung 58: Deformierter producing device produziert Vierecks-Icons in Gegenwart einer Fressfeindattrappe

Ein Verteidiger des Disjunktionsproblems könnte dagegen einwenden, warum nun nicht die Abbildungsregel *kleines, rötliches oder grosses, blaues Objekt – Viereck* gelten sollte. Das Disjunktionsproblem also nach wie vor besteht. Dieser Einwand überzeugt m.E. nicht, weil es normalerweise (statistisch verstanden) nicht zutrifft, dass ein grosses, blaues Objekt ein Vierecks-Icon bewirkt, vorausgesetzt die spezifische Deformierung ist nicht häufiger als 50%.

Mit Hilfe der Erläuterung der Entstehung von Missrepräsentation aufgrund defekter Funktionsträger kann zudem gezeigt werden, dass diese Erläuterung *nicht* darauf beruht sich am Verhalten des konsumierenden Subsystem zu orientieren, sondern auf den direkten Eigenfunktionen der Repräsentationen. Der Eindruck, dass sich Millikans Theorie bei der Gehaltsbestimmung ausschliesslich am konsumierenden Subsystem orientiert, kann – abgesehen von etlichen Äußerungen Millikans – vor allem im Kontext des Disjunktionsproblems entstehen (Vgl. Abschnitt 2.1.2). Dies kann wiederum an unserem Beispiel erläutert werden. Der intakte Organismus begegnet einem grossen, blauen aber ungefährlichen Objekt und ergreift daraufhin die Flucht. Man könnte nun für die Abbildungsregel: *Nahrung – Viereck*; *Gefahr – Dreieck* plädieren und behaupten, dass die Flucht vor einem ungefährlichen Objekt keinen Selektionsvorteil bietet und daher die „Dreiecks-Repräsentation“ eine Missrepräsentation ist, was m.E. auch Millikans Position entspräche. Dadurch kann der Eindruck entstehen, dass wir das, was die Repräsentationen des Organismus repräsentieren, durch dessen Verhalten bestimmen. Flieht er, dann wurde „Gefahr“ repräsentiert. Dieser Eindruck täuscht jedoch. Es kommt, wie bereits erwähnt, vor allem auf die Abbildungsregel an. D.h. entscheidend für die Frage, ob nach Millikan eine

Missrepräsentation vorliegt, ist, ob die „Form“ der Repräsentation und die Umweltbedingung mit der Abbildungsregel überstimmen. Dies wird deutlich, wenn wir einen Organismus mit einem spezifisch deformierten *consuming device* anschauen (siehe Abbildung 59).

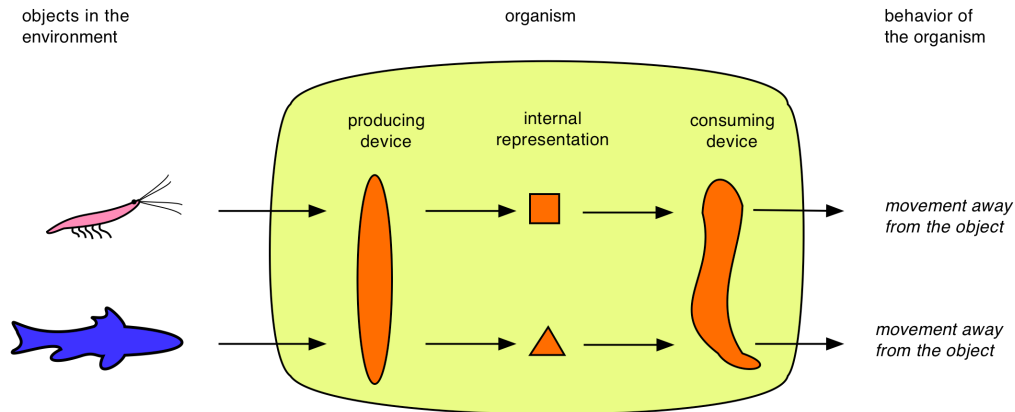


Abbildung 59: Producing device produziert Vierecks-Icon in Gegenwart von Nahrung. Deformierter consuming device reagiert auf das Vierecks-Icon aber mit Fluchtbewegung

Angenommen dieser Organismus begegnet einem kleinen, rötlichen und nahrhaften Objekt, dann reagiert er darauf aufgrund des defekten *consuming devices* mit Flucht. Würden wir uns nur am Verhalten orientieren, dann müssten wir hier von einer Missrepräsentation sprechen. Eine solche liegt aber nach Millikans Theorie nicht vor, denn die „Vierecks-Repräsentation“ in Anwesenheit eines kleinen, rötlichen Objektes entspricht der Abbildungsregel. Dabei muss man aber ein unabhängiges Kriterium für das Vorliegen der verschiedenen Repräsentationstypen und der Abbildungsregeln angeben können. Solche Kriterien könnten auf neurophysiologischen Untersuchungen und Langzeitstudien¹³⁷ basieren, wobei die Repräsentationstypen (Dreieck/Viereck) bestimmten neuronalen Erregungsmustern entsprechen könnten.

Insgesamt können wir also festhalten, wenn Millikans Annahme von missgebildeten REF-Mitgliedern gerechtfertigt ist, dann bieten diese im Rahmen ihrer Theorie eine echte Lösung für das Problem der Missrepräsentation.

¹³⁷ Wir hatten ja angenommen, dass wir bezüglich des fiktiven Wasserlebewesen nahezu alles wissen und dokumentiert haben.

Teil 3:

Bewertung von Millikans Lösung

3.1 Vorgehen

Nachdem wir Millikans Lösungsansätze für das Problem der Missrepräsentation ausführlich dargestellt haben, wollen wir nun die Frage beantworten, ob die Ansätze tatsächlich das Problem lösen können. Hierzu werden wir die Texte folgenden acht Autoren, die relevante Einwände gegen Millikans Lösungsansätze enthalten, eingehend behandeln und deren Überzeugungskraft kritisch prüfen:

1. Fodor (1990)
2. Pietroski (1992)
3. Keil (1993)
4. Davies (1994)
5. Perlman (2000)
6. Mendola (2008)
7. Stegmann (2009)
8. Nanay (2010, 2012)

Die Auswahl der Texte basiert nicht auf einer umfassenden Recherche aller Texte, die sich kritisch mit Millikans Theorie und ihren Lösungsansätzen des Problems der Missrepräsentation auseinandersetzen. Die Auswahl beansprucht also keine Vollständigkeit. Die ausgewählten Texte decken aber inhaltlich als auch zeitlich ein gewisses Spektrum ab. Millikans Lösungsansätze werden in Frage gestellt, indem:

- die Einhaltung der naturalistischen Ansprüche bezweifelt wird (Keil 1993, Nanay 2010, 2012),
- Millikans Gehaltsbestimmung als verfehlt kritisiert wird (Fodor 1990, Pietroski 1992, Perlman 2000, Mendola 2008),
- die Vereinbarkeit mit biologischen Tatsachen bestritten wird (Davies 1994, Stegmann 2009).

Ich werde die Texte hier in chronologischer Reihenfolge behandeln, da einige Texte aufeinander Bezug nehmen. Die Behandlung der Texte wird dabei immer nach demselben Muster erfolgen:

- 1) Kurze Erläuterung, worin der Einwand besteht und welche Relevanz er für Millikans Lösungsansätze hat
- 2) Rekonstruktion des Arguments, bestehend aus einer Darstellung des Arguments in Standardform (siehe Abschnitt 1.5) und einer Erläuterung der Prämissen
- 3) Diskussion des Arguments, kritisches Hinterfragen der Prämissen unter Berücksichtigung der Reaktionen von Millikan und anderen Autoren auf den Einwand

Im Anschluss an die Untersuchung der acht Einwände werde ich noch einen eigenen Einwand präsentieren.

3.2 Einzelne Einwände

3.2.1 Fodor: Das Disjunktionsproblem

In diesem Abschnitt geht es darum, einen Einwand von Jerry Fodor (1990) gegen Millikans Theorie des Gehalts zu rekonstruieren und anschließend kritisch zu prüfen. Fodors Einwand ist einer der zentralen Streitpunkte in der Debatte um das Problem der Missrepräsentation (Vgl. Detel 2001a; Neander 2012; Rakova 2006), insofern bedarf die Diskussion von Fodors Kritik im Rahmen dieser Arbeit keiner weiteren Erläuterung.

Fodor (1990) wirft teleosemantischen Theorien allgemein und Millikans Theorie im Besonderen vor, dass sie das Disjunktionsproblem entgegen dem ersten Schein doch nicht lösen können. Kurz zur Erinnerung (Vgl. Abschnitt 1.3.2): Das Disjunktionsproblem besteht bei einfachen kausalen Theorien, die besagen, dass der Gehalt einer Repräsentation immer die Ursache der Repräsentation ist: „Hund“ bedeutet Hund, weil Hunde gesetzmässig „Hund“-Repräsentationen verursachen (Vgl. Fodor 1990, 59). Nun gibt es zwei Möglichkeiten:

1. „Hund“ wird nur durch Hunde verursacht. Daraus folgt aber, dass „Hund“ nicht falsch werden kann. Mit anderen Worten: Missrepräsentation ist unmöglich.
2. „Hund“ wird auch durch Nicht-Hunde (z.B. Katzen bei Nacht) verursacht. Aufgrund unserer Bestimmung des Gehalts durch die Kausalrelation folgt daraus, dass der Gehalt von „Hund“ nun eine Disjunktion ist, nämlich *Hunde oder Katzen-bei-Nacht*. D.h. die Extension von „Hund“ sind Hunde und Katzen bei Nacht. Daraus wiederum folgt, dass eine „Hund“-Repräsentation, die von einer Katze bei Nacht verursacht wurde, wahr ist. Mit anderen Worten: Missrepräsentation ist unmöglich (Vgl. Fodor 1990, 59).

Das Disjunktionsproblem besteht also darin, dass zwischen den Disjunktionsgliedern nicht unterschieden werden kann (da beide potentielle Ursachen sind), weswegen Missrepräsentationen unmöglich sind.

Der teleosemantische Lösungsansatz besteht nun darin, die Funktion der Repräsentation ins Spiel zu bringen: Die „Hund“-Repräsentation hat die Funktion Hunde zu repräsentieren und nicht Katzen-bei-Nacht, deswegen ist „Hund“ verursacht durch eine Katze bei Nacht eine Missrepräsentation (Vgl. Abschnitt 1.3.4.1 für eine ausführlichere Darstellung). Dieser Lösungsansatz ist aber unbefriedigend, solange man die Funktion der Repräsentation nicht eindeutig bestimmen kann (Vgl. Abschnitt 1.3.4.2 für eine ausführlichere Darstellung)¹³⁸. Befürworter teleosemantischer Ansätze, wie Millikan, Neander und Papineau, vertreten deshalb explizit einen etiologischen Funktionsbegriff und wir haben gesehen, wie gründlich dieser Funktionsbegriff bei Millikan ausgearbeitet ist (Vgl. Abschnitte 2.1.1.3.1 - 2.1.1.3.7). Dennoch glaubt Fodor, dass auch dieser evolutionsbiologisch fundierte Funktionsbegriff unbestimmt ist. Der Grund hierfür liegt nach Fodor in der Tatsache, dass die Natürliche Selektion extensional in Folgendem Sinne ist:

"if it is adaptive for an organism, O, to do something, M, in the presence of environmental feature, F, and F is reliably co-extensive with another feature, G, then it is equally adaptive for O to do M in the presence of G." (Neander 2012)

¹³⁸ Wäre es die Funktion der „Hund“-Repräsentation Hunde oder Katzen-bei-Nacht zu repräsentieren ist „Hund“ verursacht durch eine Katze bei Nacht keine Missrepräsentation.

Daraus folgt nach Fodor, dass teleosemantische Theorien nicht zwischen zwei Gehaltskandidaten unterscheiden können, die co-extensional in der Umwelt vorkommen, in der der Organismus evolviert ist (Vgl. Neander 2012). Millikans Theorie ist dabei eines der Hauptangriffsziele von Fodors Kritik¹³⁹.

Rekonstruktion von Fodors Argument

Ich betrachte Fodors Einwand hier als direkt auf Millikans Theorie bezogen. Dementsprechend kann man Fodors Argument wie folgt rekonstruieren:

1. Naturalistische Theorien des Gehalts sind unbefriedigend, wenn sie keine Missrepräsentationen zulassen.
 2. Wenn Missrepräsentation in einer naturalistischen Theorie des Gehalts möglich sein soll, dann muss diese Theorie das Disjunktionsproblem lösen.
 3. Millikans Biosmantik ist eine naturalistische Theorie des Gehalts und ihr Lösungsansatz des Disjunktionsproblems basiert auf einem evolutionsbiologisch fundierten Funktionsbegriff.
 4. Millikans evolutionsbiologisch fundierter Funktionsbegriff ist unbestimmt, so dass das Disjunktionsproblem (nur verschoben aber) nicht gelöst wird.
-
5. Millikans Theorie ist unbefriedigend.

Nachdem wir die Struktur des Argumentes rekonstruiert haben, wollen wir nun die einzelnen Prämissen genauer anschauen.

Die *erste Prämisse* gibt die für diese Arbeit zentrale Grundannahme naturalistischer Theorien des Gehalts wieder, die besagt, dass die Theorie so konstruiert sein muss, dass Gehaltsträger auch falsch bzw. inkorrekt sein können (Vgl. Abschnitt 1.2.5).

Der *zweiten Prämisse* zufolge besteht ein unmittelbarer Zusammenhang zwischen dem Disjunktionsproblem und Problem der Missrepräsentation: Eine Lösung des Ersteren soll eine notwendige Bedingung für die Lösung des Letzteren sein.

„Indeed, it [that no token of a symbol can be false] is so not satisfactory that the question whether a naturalistic semantics is possible has recently come to be viewed as identical in practice to the question whether the disjunction problem can be solved within a naturalistic framework.“ (Fodor 1990, 60)

Auch wenn Fodor hier von der Ansicht Dritter spricht, ist es meines Erachtens klar, dass auch er dieser Ansicht ist.

Die *dritte Prämisse* bedarf ebenfalls keiner grossen Erläuterung. Wir haben Millikans Funktionsauffassung sehr ausführlich in den Abschnitten 2.1.1.3.1 - 2.1.1.3.7 behandelt. Hierbei sollte deutlich geworden sein, dass Millikans Funktionsauffassung auf evolutionsbiologischen Konzepten insbesondere dem der Natürlichen Selektion beruht. Der Rückgriff auf die Natürliche Selektion bietet sich aufgrund der naturalistischen Ansprüche geradezu an.

Fodor verweist auch darauf, dass eine Funktionsauffassung, die auf den Absichten eines Designers beruht, grundsätzlich eine Alternative wäre das Disjunktionsproblem zu lösen.

¹³⁹ Fodor (1990) kritisiert zudem Millikans Auffassung von Wünschen (desires). Auf diese Kritik werde ich hier nicht eingehen, denn sie ist grundlegend mit Millikans höheren Ansprüchen an ihre Theorie verbunden, die in dieser Arbeit keine Rolle spielen (Vgl. aber Detel 2001a, 610 zu Fodors diesbezüglicher Kritik).

„So there is no disjunction problem for derived intentionality. Where we have things whose states have derived intentionality (the intentionality of all the artifacts that Granny’s made so far, by the way) we can construe very fine distinctions among the contents of their states. That’s because we can construe very fine distinctions among the contents of *our* states, and derived intentionality is intentionality that’s derived *from us*.“ (Fodor 1990, 78)

Diese Alternative kommt aber für eine Theorie mit naturalistischen Ansprüchen, wie die Biosemantik, nicht in Frage, weil dabei Absichten vorausgesetzt werden müssen (Vgl. dazu Abschnitt 1.2.3).

Die *vierte Prämisse* ist für Fodors Argument zentral und bedarf einiger Erläuterungen. Zunächst müssen wir verstehen, inwiefern die Unbestimmtheit der Funktion das Disjunktionsproblem nur verschiebt und nicht löst, denn dies ist der Kern von Fodors Einwand:

„In the first place – contrary to advertisements that you may have seen – the teleological story about intentionality does not solve the disjunction problem. The reason it doesn’t is that teleological notions, insofar as they are themselves naturalistic, always have a problem about indeterminacy just where intentionality has its problem about disjunction.“ (Fodor 1990, 70)

Die Grundidee haben wir oben schon kurz angedeutet und wir haben sie auch schon ausführlicher bei Dretskes Beispiel von magnetotaktischen Bakterien kennengelernt (Vgl. Abschnitt 1.3.4.2). Dennoch ist es hilfreich sich das Problem an dem Beispiel klar zu machen, an dem es am meisten diskutiert wurde: Das Fliegenfangverhalten des Frosches (Vgl. auch Abschnitt 2.2.2.2).

Frösche reagieren auf kleine, dunkle, sich schnell bewegende Objekte in ihrem Sichtfeld mit einem Schnappverhalten (Auswerfen der Zunge) nach diesen Objekten. Bei diesen Objekten handelt es sich in der natürlichen Umwelt des Frosches meist um Fliegen. Neuronal gesehen passiert dabei sehr grob gesagt Folgendes: Das Objekt im Sichtfeld des Frosches löst im visuellen System des Frosch ein spezielles Nervensignal aus und dieses wiederum bewirkt das Auswerfen der Zunge nach dem Objekt. Die Frage ist nun welchen Gehalt das Nervensignal hat. Für Viele, wie Israel, den Fodor stellvertretend kritisiert (Fodor 1990, 71), steht fest, der Gehalt ist „Fliege“ und nicht „kleines, dunkles, sich schnell bewegendes Objekt“, weil der Frosch hinter Fliegen her ist und normalerweise kleine, dunkle, sich schnell bewegende Objekte im Sichtfeld des Frosches Fliegen sind. Für Israel, wie viele andere, ist also die Funktion des Mechanismus, der das Nervensignal produziert, dieses Signal in Gegenwart von vorbeifliegenden Fliegen zu produzieren. Dementsprechend gilt das spezifische Nervensignal als eine Missrepräsentation, wenn es in Gegenwart eines vorbeifliegenden Plastikpellets (bee bee) produziert wird.

Fodor wendet nun gegen dieses Vorgehen ein, dass man die Funktion des Mechanismus’, der das spezifische Nervensignal produziert, ohne Weiteres auch anders hätte beschreiben können: Die Funktion des Mechanismus’ ist es das spezifische Nervensignal in Gegenwart von vorbeifliegenden kleinen, schwarzen Objekten zu produzieren. Unter dieser Beschreibung ist das spezifische Nervensignal dann keine Missrepräsentation, wenn es in Gegenwart eines vorbeifliegenden Plastikpellets (bee bee) produziert wird.

„The trouble is, however, that this doesn’t *solve* the disjunction problem; it just *begs* it. For, though you *can* describe the teleology of frog’s snap-guidance mechanism the way Israel wants you to – in Normal circumstances, it resonates to flies; so its function is to resonate to flies; so its intentional content is *about* flies – there is precisely nothing to stop you from telling the story in quite a different way. On the alternative account, what the neural mechanism in question is designed to respond to is little ambient black things. [...] Hence, a frog is responding *Normally* when, for example, it snaps at a little

ambient black thing that is in fact *not* a fly but a bee-bee that happens to be passing through.“ (Fodor 1990, 71-72)

Fodors Punkt ist also, dass es zu jeder Funktionsbeschreibung, die bestimmte Repräsentationen als Missrepräsentationen bestimmt, wodurch Missrepräsentation allgemein möglich wäre, eine gleichwertige alternative Funktionsbeschreibung gibt, die dieselben Repräsentationen als korrekt bestimmt, wodurch Missrepräsentation unmöglich zu sein scheint. Mit anderen Worten: Der teleosematische Lösungsansatz des Disjunktionsproblems beruht letztlich auf einer willkürlichen Entscheidung und ist daher keine echte Lösung. Für Fodors Argumentation ist es sehr wichtig, dass die Alternativen gleichwertig sind, andernfalls hätten die Teleosemantiker ja einen guten Grund eine der beiden Alternativen abzulehnen und die Funktionsbeschreibung wäre nicht mehr willkürlich.

Der Grund dafür, dass es nach Fodor immer gleichwertige Alternativen gibt, ist, wie oben bereits erwähnt, die Tatsache, dass die Natürliche Selektion extensional in Folgendem Sinne ist:

„if it is adaptive for an organism, O, to do something, M, in the presence of environmental feature, F, and F is reliably co-extensive with another feature, G, then it is equally adaptive for O to do M in the presence of G.“ (Neander, StanEnc)

Deswegen nützt Fodor zufolge der Bezug auf Darwins Prinzip der Natürlichen Selektion nicht im Geringsten um eine Wahl zwischen bestimmten alternativen Funktionsbeschreibungen zu motivieren.

„It bears emphasis that Darwin doesn’t care which of these ways you tell the teleological story“ (Fodor 1990, 72)

Betrachten wir nun wie sich die zuvor beschriebene Extensionalität der Natürlichen Selektion in unserem Froschbeispiel auswirkt. Hierbei lohnt es sich, wie Sterelny (1990, 126), zwei Varianten zu unterscheiden, auf die sich Fodor beruft.¹⁴⁰

Variante A

Wenn es keine bee-bees aber Fliegen gibt, dann ist das Schnappen (M) des Frosches (O) nach Fliegen (F) ebenso adaptiv wie ein Schnappen nach Fliegen-oder-bee-bees (G).¹⁴¹ Da es ex hypothesi keine bee-bees gibt, haben „Fliegen“ und „Fliegen-oder-bee-bees“ die gleiche Extension. „Fliegen“ und „Fliegen-oder-bee-bees“ sind ko-extensional.

„[...] it’s equally OK with Darwin which way you describe the intentional objects of fly snaps, so long as it is reliable [...] that all the local flies-or-bees are flies.“ (Fodor 1990, 73)

Variante B

Wenn alle schwarzen Punkte in der Umgebung Fliegen sind und alle Fliegen schwarze Punkte in der Umgebung, dann ist das Schnappen (M) des Frosches (O) nach Fliegen (F) ebenso adaptiv wie ein Schnappen nach schwarzen Punkten in der Umgebung (G).

„Darwin doesn’t care how you describe the intentional object of frog snaps. All that matters for selection is how many flies the frog manages to ingest in consequence of its snappings, and this number comes out exactly the same whether one describes the function of snap-guidance mechanisms with respect to a world that is populated by

¹⁴⁰ Warum dies lohnend ist, wird später deutlich werden.

¹⁴¹ Fodor spricht auch von „fleebees“ als primitive Repräsentation mit disjunktiver Extension. Damit will Fodor folgendem Einwand entgegen: Der Frosch benötigt für eine „Fliegen-oder-bee-bees“-Repräsentation eine separate „bee-bees“-Repräsentation. Letztere ist evolutionsbiologisch unnütz, also hat der Frosch sie nicht und folglich auch keine „Fliegen-oder-bee-bees“-Repräsentation. Es bleibt nur die Option „Fliege“ (Vgl. Fodor 1990, 73-74).

flies that are, de facto, ambient black dot, or with respect to a world that is populated by ambient black dot that are, de facto, flies.“ (Fodor 1990, 72-73)

Es bringt also nichts, so Fodors Fazit, sich in irgendeiner Weise Hilfe von Darwin bei der Lösung des Disjunktionsproblem zu erwarten, denn:

„Darwin cares how many flies you eat, but not what description you eat them under.“
(Fodor 1990, 73)

Diskussion von Fodors Argument

Fodors Einwand wurde viel diskutiert und kritisiert¹⁴². Ich werde hier nicht jeden Gegeneinwand berücksichtigen, sondern konzentriere mich auf Millikans Erwiderung (1991) und Neanders sehr erhellende Beiträge (Neander 1995; 2012).

Das Argument wurde so rekonstruiert, dass der Schluss aus den Prämissen folgt. Es bleibt also kritisch zu prüfen, ob auch die Prämissen wahr sind.

Die *erste Prämisse*, die besagt, dass naturalistische Theorien des Gehalts Missrepräsentationen zulassen müssen, setze ich selbst in dieser Arbeit voraus, und werde sie daher nicht in Frage stellen.

Die *zweite Prämisse* besteht in der Annahme, dass die Lösung des Disjunktionsproblems eine notwendige Bedingung für die Lösung des Problems der Missrepräsentation ist. Diese Annahme ist sehr fragwürdig. Wie wir in den Abschnitten 2.2.1 – 2.2.3 klar herausarbeiten konnten, bietet Millikans Theorie zwei Lösungen für das Problem der Missrepräsentation. Insbesondere der zweite Lösungsansatz über die deformierten Funktionsträger erlaubt eine Lösung des Problems der Missrepräsentation, ohne zugleich das Disjunktionsproblem zu lösen. Im Abschnitt 2.2.3.2 hatten wir dies bereits an unserem fiktiven Wasserorganismus-Beispiel dargelegt. Es mag aber hilfreich sein diesen Punkt aufgrund seiner Wichtigkeit noch einmal im Kontext des Froschbeispiels zu erläutern. Beim Disjunktionsproblem gehen wir im vornherein von einem gesunden bzw. intakten Frosch aus und fragen uns dann, welchen Gehalt das spezifische Nervensignal hat, das in Gegenwart von dunklen, kleinen, bewegten Objekten zum Ausschnellen der Zunge in Richtung dieser Objekte führt. Ist der Gehalt:

1. Fliege *oder*
2. dunkles, kleines, bewegtes Objekt *oder*
3. Fliege oder Bee-Bee (fleebee)?

Das ist das Disjunktionsproblem. Nehmen wir nun an, der Frosch hat durch irgendwelche Ereignisse einen neuronalen Schaden erlitten, der insbesondere den sensorischen Apparat betrifft. Der Schaden ist derart, dass der sensorische Apparat in Gegenwart von hellen, grossen, unbewegten Objekten das spezifische Nervensignal produziert und daraufhin nach diesen Objekten schnappt. Das spezifische Nervensignal sollte nun in jedem Fall als Missrepräsentation gelten, denn es ist völlig egal für welchen der drei Gehalte (Fliege/dunkles, kleines, bewegtes Objekt/Fliege oder Bee-Bee (fleebee)) wir schließlich eine Begründung finden würden, in keinem Fall stimmt Gehalt und Wirklichkeit überein. Mit anderen Worten: Das Disjunktionsproblem bleibt ungelöst und trotzdem sind Missrepräsentationen möglich. Dieser Zug ist teleosemantischen Theorien mit einem etiologischen Funktionsbegriff, wie Millikans Theorie, vorbehalten.

Millikan selbst bringt diesen Punkt aber in ihrer Erwiderung auf Fodor nicht an (Vgl. Millikan 1991). Sie beruft sich meines Wissens auch an keiner anderen Stelle ihres Werkes auf diesen Punkt. Dies mag damit zusammenhängen, dass Millikan aufgrund ihres Anspruches (Vgl.

¹⁴² Für einen Überblick siehe Neander (2012) und Detel (2001a).

Abschnitt 1.4.3) glaubt, dass Disjunktionsproblem ohnehin lösen zu müssen. So soll Millikans Theorie nach ihrem eigenen Anspruch auch menschliche Überzeugungen erfassen. Diese können aber auch falsch sein, ohne dass ein neuronaler Schaden vorliegt, also benötigt Millikan eine Möglichkeit der Missrepräsentation bei gesunden/intakten Organismen.

Karen Neander ist meines Wissens die Erste und Einzige, die die oben dargelegte Argumentationslinie verfolgt hat.

„Perhaps surprisingly, my aim is not to provide a solution to what has become known as the ‚disjunction problem‘. Although this problem needs solving, I argue that we don’t need to solve it in order to solve the misrepresentation problem.“ (Neander 1995, 110)

Vor allem die Annahme, dass der Frosch gar nicht missrepräsentieren könne, wenn nicht das spezifische Nervensignal in Reaktion auf ein Bee-Bee als Missrepräsentation gilt, wird von Neander zu Recht in Frage gestellt. Ein Frosch, der aufgrund einer physiologischen Beeinträchtigung das spezifische Nervensignal in Gegenwart einer Schnecke¹⁴³ produziert, hat eine Missrepräsentation auch wenn, wie nach Neanders Auffassung, das Nervensignal in Bezug auf ein Bee-Bee keine Missrepräsentation ist.

„[...] if the frog R-tokens at anything which reflects onto its retina a pattern that falls outside of the specified parameters, then it misrepresents. The images cast by snails, for example, fall outside these parameters [...], so R-tokening in response to a snail is a misrepresentation. [...] It’s true that they can’t/don’t when they are functioning properly, but they don’t always function properly. A sick frog might R-token at a snail if it was dysfunctional in the right way. Damaging the frog’s neurology, interfering in its embryological development, tinkering with its genes, giving it a virus, all of those could introduce malfunction and error.“ (Neander 1995, 131)

Neander benötigt diese Möglichkeit der Missrepräsentation dringend, denn nach ihrer Auffassung hat das spezifische Nervensignal, das in Gegenwart eines Bee-Bee produziert wurde, den Gehalt kleines, schwarzes Objekt und folglich ist es keine Missrepräsentation¹⁴⁴. Deswegen behauptet Neander im obigen Zitat auch, dass der gesunde/intakte Frosch nicht missrepräsentieren kann. Soviel zur zweiten Prämisse.

Die *dritte Prämisse* kann als unkontrovers gelten. Die evolutionsbiologische Fundierung von Millikans Funktionsbegriff ist abgesehen von wenigen Einwänden (Vgl. Abschnitt 2.1.1.3.4 und 3.2.4) kaum in Frage gestellt worden. Da Millikan selbst Fodors Annahme, ihr Funktionsbegriffs sei evolutionsbiologisch fundiert, akzeptiert, wollen wir dies auch tun¹⁴⁵.

Kommen wir zur *vierten Prämisse*. Fodor nimmt hier an, dass das Disjunktionsproblem nicht unter Rückgriff auf einen evolutionsbiologisch fundierten Funktionsbegriff gelöst werden kann, weil es immer darwinistisch gesehen gleichwertige alternative Funktionsbeschreibungen gibt, so dass jede Missrepräsentation unter der einen Funktionsbeschreibung zu einer korrekten Repräsentation unter der anderen Beschreibung wird. Dass diese gleichwertigen, alternativen Funktionsbeschreibungen existieren, liegt Fodor zufolge daran, dass die Natürliche Selektion im oben beschriebenen Sinne extensional ist. Daraus folgt für Fodor insbesondere, dass mit Hilfe von Darwin nicht zwischen Gehaltskandidaten unterschieden werden kann, die koextensional in der Umwelt des betreffenden Organismus vorkamen (Vgl. Neander 2012).

Die letzte Schlussfolgerung kann mit Hilfe von Sobers Unterscheidung zwischen *selection of* und *selection for* (Vgl. Sober 1984b, 50-51) in Frage gestellt werden. Auch Millikan verweist

¹⁴³ Während die Fliegen, Bee-Bees und andere kleine, schwarze, sich schnell bewegende Objekte die Eigenschaften *klein, schwarz, sich schnell bewegend* miteinander teilen, gilt dies für die Schnecke nicht.

¹⁴⁴ Dazu später mehr.

¹⁴⁵ Millikan moniert lediglich, dass Fodor die Komplexität ihrer Auffassung unberücksichtigt lässt. (Millikan 1991, 162).

auf diese Unterscheidung (Vgl. Millikan 1991, 159) und sie gilt mittlerweile als Standardreaktion auf Fodors Einwand (Vgl. Neander 2012). Sober macht seine Unterscheidung zwischen *selection of* und *selection for* an einem Spielzeug seines Sohnes klar. Das Spielzeug ist eine Art Sieb mit mehreren Ebenen, wobei die Ebenen zunehmend kleinere Löcher haben. Zu dem Spielzeug gibt es Kugeln unterschiedlicher Grösse. Kugeln gleicher Grösse haben dieselbe Farbe: die kleinsten Kugeln sind schwarz, die Mittleren grau und die Grossen weiß. Gibt man die Kugeln in das Sieb, dann passieren am Ende nur die kleinen, schwarzen Kugeln das gesamte Sieb. Das Sieb selektiert also schwarze und kleine Kugeln (*selection of*), aber es selektiert eindeutig nur nach der Grösse, also nur nach kleinen Kugeln (*selection for*):

„It is gratifying to find these biological ideas already enshrined in the ordinary meanings of 'selection for' and 'selection of'. My young son has a toy which takes all the mystery out of this distinction. Plastic discs with circles cut out of them are stacked with spaces in between in a closed cylinder. The top-most disc contains very big holes, and the holes decrease in size as one moves down from disc to disc. At the top of the cylinder are found balls of different sizes. A good shaking will distribute the balls to their respective levels. The smallest balls end up arrayed at the bottom. The next smallest sized balls settle at the next level up, and so on. It happens that the balls of the same size also happen to have the same color. Shaking sends the black balls to the bottom, the pink ones to the next level up, and so on. The whole cylinder (plus paternally administered shaking) is a selection machine. The device selects for small balls (these are the ones which pass to the bottom). It does not select for black balls (even though these are the ones which pass to the bottom). But when we ask after a shaking what was selected, it is equally correct to say that the black balls were selected and that the small ones were. 'Selection for' focuses on causes; 'selection of' picks out effects.“ (Sober 1984b, 50-51)

Selection for bezieht sich also im Gegensatz zu *selection of* auf die für den Selektionsprozess kausal relevanten Eigenschaften. Damit gibt es aber eine prinzipielle Möglichkeit auch bei koextensionalen Eigenschaften unter Berufung auf die Natürliche Selektion zu entscheiden, auf welche der beiden Eigenschaften die interne Repräsentation referiert. Dies gilt zumindest in den Fällen, in denen nur eine der beiden Eigenschaften für den Selektionsprozess kausal relevant ist.

Angenommen Fliegen und nur Fliegen senden eine Strahlung der Frequenz X aus, d.h. „Fliege“ und „Strahlungsquelle der Frequenz X“ sind koextensional. Weiterhin angenommen die Eigenschaft des Detektor-Schnapp-Mechanismus' de facto auf Strahlungsquellen der Frequenz X zu reagieren, ist völlig irrelevant für dessen Selektion gewesen. D.h. weder detektiert er das Objekt aufgrund der spezifischen Strahlung, noch hat die spezifische Strahlung des Objektes irgendeinen nennenswerten positiven Effekt auf den Organismus. Im Gegensatz dazu ist die Eigenschaft eine Fliege zu sein relevant für den Selektionsprozess des Detektor-Schnapp-Mechanismus¹⁴⁶. Einerseits detektiert der Frosch die Fliege aufgrund der typischen äusserlichen Eigenschaften: Grösse, Farbe, Bewegung (kleines, dunkles, sich schnell bewegendes Objekt), andererseits trägt die typische chemische Zusammensetzung der Fliege dazu bei den Frosch zu ernähren. Wenn wir uns also von einem strikt darwinistischen Standpunkt aus entscheiden müssten, worauf die interne Repräsentation des Frosches reflektiert, dann fällt die Wahl klar auf Fliege und nicht auf Strahlungsquelle der Frequenz X.

Analog können wir jetzt begründen, warum „Fly“ und „Fly or bee-bee (fleebee)“ keine gleichwertigen Alternativen sind. Da es nach Fodors eigener Annahme keine Bee-Bees gibt bzw. gegeben hat, können diese auch keinen kausalen Einfluss auf den Selektionsprozess

¹⁴⁶ Man müsste eigentlich genauer sagen, dass es bestimmte Eigenschaften sind, die mit dem Fliegesein verbunden sind. Wir werden später bei der Diskussion von Agars Vorschlag darauf zurückkommen.

gehabt haben und aufgrund dieser kausalen Irrelevanz von Bee-Bees sollte die Wahl klar auf Fliegen fallen. In diesem Sinne argumentiert auch Millikan dafür, dass Flee-Bees zu fangen aufgrund der kausalen Irrelevanz von Bee-Bees keine Funktion des Detektor-Schnapp-Mechanismus sein kann.

„Disjunctive properties can enter into the premises of inferences from knowledge of earlier situations to knowledge of later ones, but this kind of deductive inference from past conditions to future ones often has nothing to do with causal explanation. That frog’s ancestors caught flies-or-bee-bees and there were no bee-bees, is no more part of a causal explanation for the proliferation of frogs, than that I caught the measles or the flu and I didn’t catch the flu, is part of a causal explanation of my red rash. Hence the frogs fly-catching mechanisms assuredly are not flee-bee catchers.“ (Millikan 1991, 160)

Neander weist zudem darauf hin, dass es für Fodors Argumentation im Prinzip unerheblich ist, dass bee-bees zumindest einige Eigenschaften (klein, schwarz, sich schnell bewegend) mit Fliegen teilen. Solange es keine Meerjungfrauen gibt, hätte die Disjunktion auch „Fliegen oder Meerjungfrauen“ lauten können¹⁴⁷. Kein Biologe würde jedoch eine solche Beschreibung akzeptieren. Damit hätten wir auch einen pragmatischen Grund gefunden die Alternative „Fly“ und „Fly or bee-bee (fleebee)“ nicht als gleichwertig zu betrachten. Darüber hinaus wird hierdurch nochmals besonders deutlich, dass Fodor die Koextensionalität für die einzige Quelle des Disjunktionsproblems hält. Dass dies nicht korrekt ist, werden wir unten noch sehen.

„People rarely notice that Fodor’s principle – that natural selection is extensional – generates a host of function attribution that no biologist (qua biologist) would take seriously for a moment. To illustrate the disjunction problem we might as well have said that since there were no mermaids in the environment in which the bacteria evolved, the kind de-oxygenated water was (and is) co-extensive with the kind de-oxygenated water or mermaids, so that directing the bacteria to de-oxygenated water or mermaids was equivalent to directing the bacteria to de-oxygenated water [...]. This case, unlike Dretske’s original example, has only one source, and it is the one which Fodor identifies as the source of functional indeterminacy – the description insensitivity of natural selection.“ (Neander 1995, 121)

Ein weiterer Einwand gegen die Alternative „Fly or bee-bee (fleebee)“ soll hier nur kurz erwähnt werden. Er stammt von Kim Sterenly. Sterenly wirft Fodor vor, dass er mit seinem Bezug auf Prädikate wie „fleebee“, die ganz analog wie Goodmans „grue“ funktionieren, ein Standardproblem in der Wissenschaftsphilosophie zu Unrecht als einen spezifischen Einwand gegen teleosemantische Theorien verwendet¹⁴⁸.

„I think Fodor is presenting a standard problem in philosophy of science as though it were one that especially infects teleological explanation. There is a general problem with dodgy predicates. The classic example, due to Goodman, is ‚grue‘. [...] But this problem has nothing especially to do with biological or psychological explanation.“ (Sterenly 1990, 126-127)

Es scheint also etliche Gründe zu geben, warum wir die Gleichwertigkeit der Alternativen „Fly“ und „Fly or bee-bee (fleebee)“ ablehnen sollten.

Wie steht es nun aber mit der zweiten Variante von Fodor, „Fly“ oder „black dot“? Sobers Unterscheidung hilft hier nicht mehr weiter, denn beide Alternativen beziehen sich auf

¹⁴⁷ Neander bezieht sich allerdings auf Dretskes Beispiel der Magnetotaxis von Bakterien (siehe Zitat unten).

¹⁴⁸ Diesem Einwand stimmen auch Goode und Griffith zu, wenngleich sie dessen Begründung ablehnen (Vgl. Goode & Griffith 1995, 104).

Eigenschaften, die kausal relevant für den Selektionsprozess des Detektor-Schnapp-Mechanismus waren. Anders als bei unserem Beispiel mit der Strahlungsquelle der Frequenz X nutzt der Forsch die Eigenschaften schwarz und klein (black dot) um das betreffende Objekt zu detektieren, deswegen sind diese Eigenschaften auch kausal relevant für den Selektionsprozess.

„The distinction [Sober’s, AW] is supposed to be between properties that were causally efficacious in a selection process (those that were selected for) and those that were not (the free riders) [...] but several properties can be causally efficacious in the same selection process: in the case of the frog, one as the benefit (being frog-food) and another as the stimulus (being small dark and moving). If so, Sober’s distinction doesn’t cut between them.“ (Neander 1995, 128)

Man könnte auch sagen, dass der Detektor-Schnapp-Mechanismus zugleich mehrere Funktionen hat:

- a) auf kleine, schwarze, sich schnell bewegende Objekte zu reagieren und
- b) auf Fliegen (Froschfutter) zu reagieren

Dies ist Millikans Ansicht¹⁴⁹. Millikan bemerkt zudem richtig, dass diese multiplen Funktionen nichts mit Koextensionalität zu tun haben (Vgl. Millikan 1991, 161)¹⁵⁰, weswegen Neander von vornherein zwei Unbestimmtheitsprobleme unterscheidet:

„there are two [...] indeterminacy problems. One concerns the fact that natural selection is extensional [...] and the other concerns the fact that natural selection selects traits for complex causal roles“ (Neander 2012)

Neander zufolge sind die multiplen Funktionen ein und desselben Funktionsträgers auch nicht unabhängig voneinander, sondern durch eine *by-relation* verbunden. D.h., indem die eine Funktion ausgeübt wird, wird auch die andere Funktion ausgeübt (aber nicht umgekehrt). Indem der Forsch nach kleinen, schwarzen Objekten schnappt, schnappt er zugleich nach Fliegen.

Neander erläutert diesen Punkt ausführlich an einem Beispiel. Angenommen in einer Antilopenpopulation tritt eine Mutation auf, die eine veränderte Hämoglobinstruktur zur Folge hat. Die neue Hämoglobinstruktur erlaubt eine erhöhte Sauerstoffaufnahme. Dies wiederum erlaubt es den Trägern der Mutation in grösseren Höhen zu überleben, in die ihre natürlichen Feinde nicht gelangen. All dies führt letztlich dazu, dass sich die Träger der Mutation in der Population ausbreiten. Was ist nun die Funktion der Mutation? Neanders Antwort ist: Alle diese Dinge (Vgl. Neander 1995, 114-115).

„One and the same trait can have the function of contributing to survival and reproduction, allowing the antelope to move to higher ground, increasing oxygen uptake and altering the structure of hemoglobin, because all these are correct descriptions of what the trait did which caused it to be selected.“ (Neander 1995, 115)

Schematisch lässt sich dies so darstellen:

¹⁴⁹ “And so with the magnetosome and with the frog’s detecting mechanism, they have multiple functions.” (Millikan 1991, 161)

¹⁵⁰ Es ist einfach nicht der Fall, dass alle kleinen, sich schnell bewegenden, schwarzen Objekte in der Umwelt des Frosches Fliegen sind. Es werden z.B. auch kleine Holzstücke gelegentlich vom Wind aufgewirbelt etc.

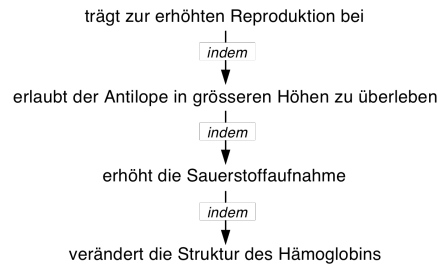


Abbildung 60: Multiple Funktionen, die durch eine *indem-Relation* (*by-relation*) verbunden sind, Antilopen-Beispiel (in Anlehnung an Neander 1995, 115).

Die *by-relation* ist also eine Relation zwischen Mechanismus bzw. Realisierung und Zweck bzw. Ergebnis und da sich Zwecke multiple realisieren lassen, ist die Relation asymmetrisch (Neander 1995, 115-116)¹⁵¹.

Das Schema lässt sich auch auf die Funktionen des Detektor-Schnapp-Mechanismus im Forsch-Beispiel anwenden:

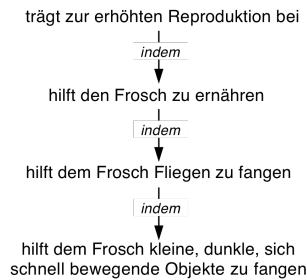


Abbildung 61: Multiple Funktionen, die durch eine *indem-Relation* (*by-relation*) verbunden sind, Frosch-Beispiel (in Anlehnung an Neander 1995, 125).

Im Folgenden gehe ich, wie Millikan und Neander, davon aus, dass ein Funktionsträger multiple Funktionen im eben beschriebenen Sinn haben kann (Vgl. auch Abschnitt 2.1.1.3.3, S. 75-77).

Was bedeutet dies nun für die Bestimmung des Gehalts?

Soweit ich sehe, gibt es zwei Möglichkeiten den Gehalt dennoch festzulegen:

- 1) Wir zeichnen eine der multiplen Funktionen als besonders aus und diese Funktion ist dann massgeblich für die Gehaltsbestimmung. Hierbei benötigt man allerdings eine gute Begründung für die Wahl, andernfalls behält Fodor bezüglich dem Bestehen gleichwertiger Funktionsbeschreibungen Recht, auch wenn seine Begründung falsch war.
- 2) Alle Faktoren, die in der Umwelt kausal für die Selektion des Funktionsträgers verantwortlich waren, gehen in den Gehalt ein.

Für die erste Möglichkeit plädieren u.a. Sterelny, Sharpiro und Neander. Während Sterelny und Sharpiro aber zugunsten einer Funktionsbeschreibung auf einer höheren Ebene argumentieren, so dass „Fliege“ oder „Fliegennahrung“ der Gehalt der internen

¹⁵¹ Auch koextensionale Beschreibungen können der soeben beschriebenen *by-relation* ähnliche Relationen generieren. Ich kann der Bundeskanzlerin zuwinken, indem ich Angela Merkel zuwinke und umgekehrt. Anders als die *by-relation* sind diese Relationen jedoch symmetrisch (Vgl. Neander 1995, 115-116). Neander Beschreibung der *by-relation* erinnert zudem stark an Anscombes Diskussion der Beschreibung von absichtlichen Handlungen (Vgl. Anscombe 1963, 37-41).

Repräsentation wäre, argumentiert Neander für Funktionsbeschreibung der unteren Ebene, so dass der Gehalt „kleines, schwarzes Objekt“ resultiert.

Betrachten wir zunächst Sterelnys Argumentation¹⁵². Diese basiert auf folgender Überlegung: Hätte sich die Form der Fliegen im Verlauf der Evolution des Frosches geändert, etwa hin zu grossen, hellen Objekten, dann wäre ein Detektor-Schnapp-Mechanismus selektiert worden, der nach wie vor auf Fliegen reagiert, aber nicht mehr auf kleine, dunkle Objekte.

„Had flies changed so that their retinal projections were not beebeelike, natural selection would tend to construct mechanisms that tracked the changing shape of flies... There is a diachronic tracking between flies and fly representors.“ (Sterelny 1990, 127)

Also – so Sterelny – ist die ausgezeichnete Funktion des Detektor-Schnapp-Mechanismus Fliegen zu fangen und nicht kleine, dunkle Objekte zu fangen und der Gehalt der inneren Repräsentation ist folglich „Fliege“ und nicht „kleines, dunkles Objekt“. Sterelny beruft sich bei seiner Argumentation ausdrücklich auf Sobers Unterscheidung zwischen *selection of* und *selection for* (Vgl. Sterelny 1990, 127).

Goode und Griffith haben Sterelnys Gebrauch von Sobers Unterscheidung als verfehlt kritisiert, weil dadurch die Beschreibungen auf verschiedenen Ebenen fälschlicherweise als konkurrierende darstellt werden.

"[...] using the selection for/of distinction across explanatory levels misrepresents complementary evolutionary narratives as competing ones." (Goode & Griffiths 1995, 105)

Das Problem mit Sterelnys Vorgehen – so Goode & Griffiths – ist vor allem, dass man nicht beim „Fliegen“-Level stehen bleiben muss. Der gleichen Logik folgend kann man dafür argumentieren, dass nicht „Fliege“, sondern „Froschfutter“ der korrekte Gehalt ist¹⁵³ und schließlich für „Fitnessverstärkendes“ anstatt „Froschfutter“. Dementsprechend wäre die einzige Funktion des Detektor-Schnapp-Mechanismus’ die Fitness des Frosches zu erhöhen, was Goode & Griffiths für absurd halten.

„If small dark moving things were not flies, there would be no selective advantage in detecting small dark moving things. Likewise, if flies were not food, then there would be no selective advantage in detecting flies. So in both cases, it seems, there is no selection for the lower-level property. We suggest that, by parity of reasoning, we can infer that there is no selection for detecting food either. For if detecting food was not fitness enhancing, then there would be no selection for it. [...] The conclusion that the sole function of state S is to increase the frog's fitness is absurd.“ (Goode & Griffiths 1995, 105)

Zudem haben Agar und Neander zu Recht darauf hingewiesen, dass für Sobers Unterscheidung die tatsächliche Selektionsgeschichte relevant ist und keine hypothetische (Vgl. Agar 1993, 9-10; Neander 1995, 128).

Sterelny (1991) bringt noch eine zweite Begründung, warum wir „Fliege“ als Gehalt der internen Repräsentation dem Gehalt „kleines, dunkles Objekt“ vorziehen sollten. Der Grund ist, dass die Beschreibung *Fliegen-Detektor*¹⁵⁴ im Gegensatz zu *Kleine-dunkle-Objekte-Detektor* speziesübergreifende Generalisierungen zulässt, weil der Detektor für Fliegen in

¹⁵² Sharpisros Argumentation ist ganz ähnlich (Vgl. Goode & Griffith 1995).

¹⁵³ Genau dies macht Shapiro (1992).

¹⁵⁴ Sterelny bezieht sich hier nur auf den Detektor-Teil des Detektor-Schnapp-Mechanismus. Seine Argumentation ließe sich aber problemlos auf den gesamten Detektor-Schnapp-Mechanismus übertragen.

verschiedenen Spezies unterschiedlich realisiert sein kann (z.B. in Fröschen und Kröten¹⁵⁵). Dennoch gibt es für die Detektoren der Frösche und der Kröten die gleiche evolutionsbiologische Erklärung, genau dieser Einsicht würden wir uns aber berauben, wenn wir den Detektor auf der niedrigeren Ebene als *Kleine-dunkle-Objekte-Detektor* beschreiben (Vgl. Sterelny 1990, 126).

Neander erwidert darauf, dass aus der sinnvollen Forderung nach speziesübergreifenden Generalisierungen und den daraus resultierenden Erklärungen, aber keineswegs folgt, dass der Gehalt der internen Repräsentation entsprechend dem höheren Level bestimmt werden müsse. Sie macht dies durch eine analoge Betrachtung des Verdauungssystems deutlich:

„Both sheep and humans have mammalian digestive systems, about which certain generalisations can be made, but what is normal functioning for the one need not be normal functioning for the other. We don't malfunction just because we can't digest grass. Similarly, the frogs detection device can be categorised as a prey-detector without committing us to any fine-grained judgements about biological norms, or semantic norms neither. We still can ask what it represents its prey as.“ (Neander 1995, 129)

Zwischenfazit: Die hier diskutierten Begründungen für die Auszeichnung einer Funktionsbeschreibung der höheren Ebenen können nicht so recht überzeugen.

Für den Vorzug einer Funktionsbeschreibung auf der niedrigsten Ebene¹⁵⁶ plädiert Neander. Ihre Begründung ist, dass diese Beschreibung der Funktion viel informativer ist insbesondere in Bezug auf Fehlfunktionen¹⁵⁷. Neander illustriert dies an folgendem Beispiel (Vgl. Neander 1995, 120): Der Eileiter einer Frau ist blockiert und sie ist deshalb unfruchtbar. Zwei Beschreibungen sollen zur Option stehen:

- 1) das Reproduktive System der Frau hat eine Fehlfunktion und alle Teile des Systems ebenfalls, denn keines kann die Funktion für eine Replikation der Gene zu sorgen erfüllen (Funktionsbeschreibung auf der höheren Ebene)
- 2) das Reproduktive System der Frau hat eine Fehlfunktion, aber nicht alle Teile des Systems (z.B. die Eierstöcke) haben eine Fehlfunktion, sondern nur der Eileiter, weil den Teilen spezifische Funktionen zukommen (Funktionsbeschreibung auf der niedrigsten Ebene)

Wenn wir nun erfahren, dass die Eierstöcke einer Frau eine Fehlfunktion haben. Dann erfahren wir nach der ersten Option, dass das Reproduktive System der Frau eine Fehlfunktion hat. Wir wissen aber nicht, ob die Eierstöcke die Ursache der Fehlfunktion sind. Genau dies erfahren wir aber nach der zweiten Option, die insofern informativer ist, und deshalb auch von Biologen bevorzugt wird.

„We should choose the second option because it makes talk of malfunctioning more informative than the first option does. [...] Not surprisingly, biologists do not maintain that the ovaries are malfunctioning because conception is impossible because the

¹⁵⁵ Sterelnys Beispiel (Vgl. Sterelny 1991, 126).

¹⁵⁶ Bezogen auf das Frosch- und Antilopen-Beispiel oben. Neander will keine Reduktion auch eine atomare Ebene oder dergleichen (Vgl. Neander 1995, 129).

¹⁵⁷ Neander führt drei weitere Gründe auf, die hier aber nur erwähnt werden sollen: 1. Wenn Missrepräsentation ohne Fehlfunktion auf inferentiellen Prozessen beruht und Frösche zu solchen Prozessen nicht fähig sind, dann führt die Wahl der höheren Funktionsbeschreibung dazu Fröschen Fähigkeiten zuzusprechen, die sie nicht haben und ist daher abzulehnen. 2. Wenn zum Gehalt doch *Sinn* und Referenz gehören, dann hat nur die Funktionsbeschreibung auf der niedrigen Ebene etwas dazu zu sagen, wie (Sinn) ein Referenzobjekt repräsentiert wird. 3. Die Funktionsbeschreibung auf der niedrigen Ebene ist kompatibler mit der gut etablierten uninterpretierten physiologischen Analyse nach Cummins (Neander 1995, 134-135).

Fallopian tubes are blocked. Not surprisingly, they choose the second option.“
(Neander 1995, 120)

Zwischenfazit: Neanders Vorschlag scheint ein gangbarer Weg.

Neben der Möglichkeit eine der multiplen Funktionen als besonders auszuzeichnen und diese Funktion dann als massgeblich für die Gehaltsbestimmung anzusehen, besteht als Alternative die Möglichkeit alle Faktoren, die in der Umwelt kausal für die Selektion des Funktionsträgers verantwortlich waren, als für den Gehalt relevant zu betrachten. Diese Option wird von Agar (1993) vertreten, der dabei ebenfalls von Sobers Unterscheidung ausgeht. In Bezug auf den Detektor-Schnapp-Mechanismus sind die für dessen Selektion kausal relevanten Umwelteigenschaften nach Agar etwa Folgende: *dunkel, schnell bewegend, nicht gefährlich, ..., nahrhaft*. Entsprechend ist der Gehalt der internen Repräsentation S etwa „dunkel, schnell bewegend, nicht gefährlich, ..., nahrhaft“.

„So, a teleocontent based on the [...] selective histories will be, very roughly, FAST-MOVING, NON-DANGEROUS, DARK, ... FOOD in the case of S [...]. In this way, the content[s] of S [is] identified strictly with the properties in [its] environment which did the selecting work.“ (Agar 1993, 9)

Das Zitat macht deutlich, dass es Agar um eine allgemein anwendbare Bestimmung des Gehaltes geht. Ihm liegt kaum etwas daran seinen Gehaltsvorschlag für das Detektor-Schnapp-Mechanismus-Beispiel als zwingend zu verteidigen. Sollten etwa neuere Forschungen zeigen, dass eine der genannten Eigenschaften doch irrelevant oder andere Eigenschaften relevant sind, dann ändert sich der Gehalt entsprechend. In diesem Sinne wurde Agars Vorschlag von Goode und Griffith korrigiert¹⁵⁸:

"The proper outcome of adopting Agar's principles will be to make the content of S 'small, dark and moving, nutritious, fitness enhancing fly'." (Goode & Griffiths 1995, 107)

Agars Vorschlag liefert also einen bestimmten Gehalt, der evolutionstheoretisch begründet ist. Mit seinem Vorschlag ist unmittelbar verbunden, dass man akzeptieren muss, dass der Frosch etwas repräsentieren kann, was er nicht unterscheiden kann. Der Frosch repräsentiert zwar etwas als nahrhaft, er kann aber nicht zwischen nahrhaften und nicht-nahrhaften Dingen unterscheiden. Dass er dies nicht unterscheiden kann, sieht man daran dass, er gleichermaßen nach Fliegen und bee-bees schnappt. Auf diesen Punkt werden wir bei der Behandlung von Pietroskis Einwand noch einmal zurückkommen (siehe S. 148). Zudem sind nach Agars Vorschlag Missrepräsentationen – anders als bei Neander – auch ohne Missbildungen oder Verletzungen möglich.

„The frog then genuinely misrepresents when S is tokened by a beebee, for the simple reason that beebees are not non-dangerous, fast moving food.“ (Agar 1993, 9)

Zwischenfazit: Agars Vorschlag scheint ebenfalls ein gangbarer Weg.

Welche Position vertritt nun Millikan zur Unbestimmtheit des Gehalts, die sich aus den multiplen Funktionen zu ergeben scheint?

Diese Frage ist nicht leicht zu beantworten, weil Ihre Antworten relativ kurz und m.E. nicht eindeutig sind. Anstatt einer ausführlichen Antwort verweist Millikan oft einfach auf Kapitel in Millikan (1984) mit der Bemerkung, dass dort ihre Position ausführlich dargelegt sei (Vgl. Millikan 1991, 163).

¹⁵⁸ Den von Goode & Griffiths ebenfalls bei Agar festgestellten Missbrauch von Sobers Unterscheidung kann ich hingegen nicht erkennen. Zudem ist fraglich, ob der Zusatz "fitness enhancing" eine geeignete Beschreibung für eine Eigenschaft ist, die kausal für die Selektion des Funktionsträgers verantwortlich war.

Man kann Millikan so verstehen, dass sie wie Sterelny für eine Auszeichnung der Funktionsbeschreibung auf höherer Ebene plädiert (Vgl. z.B. Wild 2010a). Man kann sie aber auch so verstehen, dass sie dieselbe Position wie Agar vertritt (Vgl. Detel 2001a, 614).

Wenn man allein danach geht, welche Gehaltsbestimmung sie im Froschbeispiel verteidigt hat, dann gehört sie eindeutig ins Lager von Sterelny. Nach Millikan ist nämlich der Gehalt des spezifischen Nervensignals im Detektor-Schnapp-Mechanismus „frog food“ und ausdrücklich nicht „black or ambient or specklike“ (Vgl. Millikan 1991, 163). Millikan begründet dies mit einer nötigen Verschiebung des Fokus' vom Produzenten des Signals hin zum Konsumenten des Signals (Vgl. Millikan 1991, 162).

„[...] VERY CRUDELY the solution to [...] the fly detector problems is this. [...] the systems that use, that respond to, the frog' fly detector signals, don't care at all whether these correspond to anything black or ambient or specklike, but only whether they correspond to frog food. [...] So the firing means frog food.“ (Millikan 1991, 163)

Dass dem Konsumenten des Signals neben dem Produzenten des Signals eine wichtige Rolle zukommt, sollte uns im Abschnitt 2.1.1.3.7 klar geworden sein. Dennoch kann Millikans Lösung nicht überzeugen. Neander stellt völlig zu Recht fest, dass Millikan in ihrer Antwort die Betonung auf die falsche Stelle ihrer Theorie legt, weil das Problem allein durch eine Fokussierung auf den Konsumenten des Signals nicht gelöst werden kann (Vgl. Neander 2012): Einer der unmittelbaren Konsumenten des Signals dürfte das Motorkontrollsystem des Frosches sein und wir können dessen Funktion als ein Orientieren und Schnappen des Frosches nach Froschfutter beschreiben, aber ebenso gut können wir es als ein Orientieren und Schnappen des Frosches nach kleinen, schwarzen Objekten in der Umgebung beschreiben (Vgl. Neander 2012).

Millikans Theorie löst das Problem vielmehr durch die Bezugnahme auf Normale Bedingungen. Dies ist auch Neander Ansicht:

„However, it does not follow that Millikan's theory leaves content indeterminate. It is Millikan's appeal to Normal conditions that does more work in disambiguating the content for her.“ (Neander 2012)

Die Normalen Bedingungen für das Funktionieren des Konsumenten entsprechen den kausalwirksamen Eigenschaften der Umwelt, die in der Vergangenheit zur Selektion des repräsentationalen Systems geführt haben. Millikans Position gleicht also eher derjenigen von Agar. Zudem hat sie daraufhingewiesen, dass sie in Millikan (1984), die Wichtigkeit von Produzenten und Konsumenten als gleichwertig betrachtet. Es kommt auf deren Kooperation an.

„As a corrective to the emphasis that others in the teleosemantic business have placed on the function of the representation producers, Millikan [...] has recently been emphasizing the devices that use or 'consume' representations. The official statement of Millikan's position, LTOBC, however, emphasises producer and consumer equally.“ (Millikan 1993b, 125)

Es ist die Kooperation zwischen Produzenten und Konsumenten, die auch bei der Selektion beider entscheidend ist. Deswegen sind die Normalbedingungen für das Funktionieren des Konsumenten auch gleich den Normalbedingungen für das Funktionieren des Produzenten. D.h. beide müssen in dieser Hinsicht als Einheit betrachtet werden. Diesen Punkt scheint auch Millikan prinzipiell zu akzeptieren, wenn sie Folgendes schreibt:

„Representation consumers are devices that have been designed [...] by a selection process to cooperate with a certain representation producer. The producer likewise has been designed to match the consumer. [...] The consumer operates of course, after the producer does, and a full explanation of how the consumer has historically managed to perform its function or functions [...] would include that the producer first performed

its function properly, and it would include an explanation of how the producer's function has historically been accomplished.“ (Millikan 1993b, 126)

Allerdings möchte Millikan den Umfang der Normalbedingungen gleich wieder einschränken und zwar auf die „most proximal Normal explanation“ für den Konsumenten, wodurch die Bedingungen, die vor dem Aktivwerden des Konsumenten kausal relevant sind, nicht im Gehalt berücksichtigt würden (Vgl. Millikan 1993b, 126). Diese Einschränkung scheint mir aber nicht begründet. Insbesondere basiert sie nicht allein auf der Selektionsgeschichte des repräsentationalen Systems, noch der des Konsumenten. Die evolutionsbiologische Grundlage bildet aber die Legitimation für Millikans naturalistische Ansprüche. In diesem Sinne kann Detel (2001a, 614) behaupten, dass Millikans Theorie eine durch Agar später präzierte Lösung des Unbestimmtheitsproblems bereits enthält.

FAZIT:

Fodors Einwand kann nicht überzeugen. Insgesamt basiert der Einwand auf der falschen Annahme, dass eine Lösung des Problems der Missrepräsentation zugleich eine Lösung des Disjunktionsproblems erfordert.

In Bezug auf die Lösung des Disjunktionsproblems zeigte sich dann im Detail, dass sich eine der beiden Varianten von Fodors Einwand, nämlich Variante A, durch Sobers Unterscheidung zwischen *selection for* und *selection of* entkräften lässt.

Auch für die zweite Variante B von Fodors Einwand konnten zwei gangbare Auswege aufgezeigt werden, diejenigen von Neander und Agar. Zudem konnte gezeigt werden, dass Millikans Theorie durchaus im Sinne von Agars Ausweg interpretiert werden kann.

3.2.2 Pietroski: Intentionality and Teleological Error

In diesem Abschnitt geht es darum, einen Einwand von Paul Pietroski gegen Millikans Theorie des Gehalts zu rekonstruieren und anschließend kritisch zu prüfen (Pietroski 1992). Der Einwand hat grosse Beachtung gefunden (Vgl. Papineau u. Macdonald 2006, 5; Wild 2010a, 357-457; Neander 2012) und wird von Vielen nach wie vor als relevant betrachtet (Vgl. Wild 2010a, 357-457; Schulte 2012, 495). Pietroskis Einwand steht insofern in unmittelbarem Zusammenhang mit dem Problem der Missrepräsentation als Pietroski dafür argumentiert, dass genau dasjenige Element von Millikans Theorie, das ihr eine Lösung des Problems erlaubt, zugleich dazu führt, dass ihre Theorie unakzeptable Konsequenzen hat (Vgl. Pietroski 1992, 275)¹⁵⁹.

Rekonstruktion von Pietroskis Einwand

Die Grundidee von Pietroskis Einwand ist es anhand eines fiktiven Beispiels zu zeigen, dass Millikans Theorie des Gehalts unhaltbare Konsequenzen hat. Der Einwand ist also der Versuch einer *reductio ad absurdum* von Millikans Theorie. Deshalb kann der Einwand recht anschaulich als eine sogenannte Annahmenwiderlegung rekonstruiert werden, die allgemein die folgende Form hat:

- „
1. Nehmen wir einmal an, T wäre wahr.
 2. Die Aussage A ist wahr.
 3. Doch aus T und A folgt logisch die Aussage K.
 4. Wäre nun T wahr, so müsste auch K wahr sein, da A ja wahr ist und K aus A zusammen mit T folgt.
 5. Doch K ist falsch.
-
6. Also muss T falsch sein.“ (Tetens 2006, 33)

Pietroskis Einwand weist zwar nicht ganz die logische Strenge einer Annahmenwiderlegung auf, er ist dieser aber dennoch sehr ähnlich:

1. Angenommen, Millikans Gehaltsbestimmung wäre korrekt.
2. Das fiktive Beispiel der Kimu¹⁶⁰ ist möglich/realistisch.
3. a) doch aus 1. und 2. folgt, dass Repräsentationen, die die Kimus¹⁶¹ haben, den Gehalt *x* haben (K1) und

¹⁵⁹ “It is important to bear in mind that the distinction between causation/discrimination on the one hand, and historical importance on the other is precisely what supports the rather nifty accounts of error provided by any teleological account of content. For recall that the general form of such explanation is the following: Some mechanism M produces B-tokens when and only when property Q is instantiated; but M is supposed to produce B-tokens when and only when property P is instantiated; so false beliefs are produced when Q but not P is instantiated. In this schema, Q depends on what can cause B-tokens; or again, it depends what property the system can discriminate in virtue of having M. P depends on the purpose of the mechanism as established by historical circumstances. So the possibility of cases in which B-tokens are supposed to indicate something other than what causes them is not an unimportant spandrel of Millikan’s view.” (Vgl. Pietroski 1992, 275).

¹⁶⁰ Die Kimu sind eine fiktive Tierart (siehe unten).

- b) daraus wiederum folgt, dass die Erklärung *y* die korrekte intentionale Erklärung des Kimu-Verhaltens¹⁶² ist (K2).
4. Wäre nun Millikans Gehaltsbestimmung korrekt, so müssten auch K1 und K2 wahr sein, da das fiktive Beispiel der Kimu möglich/realistisch ist und K1 und K2 aus dem Beispiel zusammen mit Millikans Gehaltsbestimmung folgen.
 5. Doch a) K1 und b) vor allem K2 sind vollkommen unplausibel.
-
6. Millikans Gehaltsbestimmung kann nicht korrekt sein.

Nachdem wir die Struktur des Argumentes rekonstruiert haben, wollen wir nun die einzelnen Prämissen genauer anschauen.

Zur *ersten Prämisse* muss nicht viel gesagt werden. Wir haben Millikans Auffassung der Gehaltsbestimmung ausführlich dargestellt (Vgl. Abschnitt 2.1.2). Pietroski gibt diese Auffassung hinreichend korrekt, wenn auch weniger detailreich wieder. Die weggelassenen Details bzw. Ungenauigkeiten gefährden zunächst aber nicht die Hauptlinie seines Argumentes.

Für die *zweite Prämisse* ist natürlich das fiktive Beispiel zentral, das von Pietroski wie folgt beschrieben wird:

„The kimus live near a large rocky hill. Their only predators are the snorfs, carnivores who roam past the hill each morning. Kimus used to be „color-blind“. But in virtue of a genetic mutation, one particular kimu – call him Jack – came to have an internal mechanism M that produced token of a physically specifiable state type B in the presence of certain wavelengths of light. Each morning, something red on the hilltop caused Jack to form a B-token when he looked up. And Jack (like his descendants) turned out to have a „fondness“ for red things; i.e., other things being equal, Jack would move towards the distal causes of B-tokens when such tokens were produced. So each morning, Jack trudged up the hill and thereby avoided the snorfs. Natural selection took over; and Jack’s mechanism type proliferated throughout the species. There was no other reason (e.g. detection of food) for selection in favor of having the „color mechanism“.“ (Pietroski 1992, 273)

Ich möchte Pietroski hier einfach zugestehen, dass dieses Szenario aus biologischer Sicht durchaus möglich wäre und das Beispiel damit als realistisch gelten kann.

Der erste Teil der *dritten Prämisse* (3a) besagt nun konkret, dass die mutierten Kimus in Gegenwart von roten Dingen B-tokens ausbilden und dass diese nach Millikan den Gehalt „fewer snorfs this way“ (Pietroski 1992, 273f) haben. Der zweite Teil der *dritten Prämisse* (3b) besagt, dass zudem aus dem ersten Teil der Prämisse (3a) bestimmte intentionale Erklärungen für das Verhalten der Kimus folgen. Wenn die Kimus beispielsweise einen Hügel erklettern, weil sie dort etwas Rotes gesehen haben, müsste man dies nun wie folgt erklären: Die Kimus klettern auf den Hügel, weil sie Snorfs vermeiden wollen und glauben, dass der Hügel snorffrei ist (aus 3a). Ähnliche intentionale Erklärungen müsste man immer geben, wenn die Kimus sich auf etwas Rotes zu bewegen.

„But then kimus climb the hill *because they believe* that the hill is snorf-less (and, presumably, they also want to avoid snorfs. Moreover, when kimus move toward red things on other occasions (on the flat), they are *acting on the belief* that the area in question is snorf-less.“ (Pietroski 1992, 273)

¹⁶¹ in bestimmten Situationen.

¹⁶² in bestimmten Situationen.

Auch wenn die Prämisse 3b als state of the art in den Kognitionswissenschaften gelten kann, scheint sie mir erläuterungsbedürftig, weil der Zusammenhang zwischen Gehaltsträgern und intentionalen Erklärungen zumindest nicht unmittelbar klar ist. Nach Pietroski ist es zunächst der grundlegende Sinn von intentionalen Zuständen¹⁶³, dass sie uns intentionale Erklärungen des Verhaltens von Intentionalen Systemen erlauben.

„For what are intentional states, if not those states that figure in (correct) intentional explanations?“ (Pietroski 1992, 273)

„Millikans has lost track of *theoretical* function of intentional states, viz., providing plausible explanation of the behavior of intentional systems.“ (Pietroski 1992, 278)

Zudem nimmt Pietroski wahrscheinlich an, dass ein Organismus, in dem ein B-token gebildet wurde, dadurch in einem bestimmten intentionalen Zustand ist.

Die *vierte Prämisse* bedarf keiner näheren Erläuterung, da sie lediglich die Beziehungen zwischen den Prämissen explizit macht.

Der erste Teil der *fünften Prämisse* (5a) besagt, dass es völlig unplausibel ist, dass der Gehalt von B-tokens, die Kimus in Gegenwart von roten Dingen ausbilden, „fewer snorfs this way“ ist. Der Grund hierfür ist, dass es nach Pietroski zwischen jedem Gehaltsträger und dem, wovon der Gehaltsträger handelt, zumindest irgendeine kausale Wechselbeziehung geben muss. Oder anders gesagt: Wenn es zwischen bestimmten Gehaltsträgern und bestimmten Objekten überhaupt keine irgendwie geartete kausale Beeinflussung gibt, dann können diese Gehaltsträger nicht von diesen Objekten handeln.

„One needn’t hold that beliefs are about what causes them to hold that, in the absense of any causal interaction between B-tokens and snorfs, B-token are not – indeed, cannot be – about snorfs.“ (Pietroski 1992, 277)

Warum dies so ist, wird von Pietroski nicht weiter begründet. Es ist einfach seine Intuition, die, wie er selbst meint, vielleicht das Produkt von kausalen Theorien der Referenz und Reliabilistischen Wissenstheorien ist¹⁶⁴. Die Prämisse 5a lässt sich also in Form eines Mini-Arguments darstellen:

1. Wenn es zwischen bestimmten Gehaltsträgern (Überzeugungen) und bestimmten Objekten überhaupt keine irgendwie geartete kausale Beeinflussung gibt, dann können diese Gehaltsträger nicht von diesen Objekten handeln (siehe oben).
2. Zwischen B-tokens und Snorfs gibt keine irgendwie geartete kausale Beeinflussung.
3. Also können B-tokens nicht von Snorfs handeln.

Der zweite Teil der *fünften Prämisse* (5b) besagt, dass die intentionalen Erklärungen von Kimu-Verhalten (siehe oben Prämisse 3b), die aus der Gehaltszuschreibung „fewer snorfs this way“ folgen, völlig unplausibel sind.

„But such intentional explanations of kimu-behavior are about as implausible as intentional explanations can be.“ (Pietroski 1992, 276)

Die Erklärungen sind unplausibel, weil sie Überzeugungen und Wünsche über Snorfs enthalten. Die Kimus aber keine solchen Überzeugungen haben können. Dies kann man sich nach Pietroski durch einige Überlegungen der folgenden Art klar machen: Wenn ein Kimu direkt (ohne Sichtbehinderung) vor einem Snorf steht und nichts Rotes in der Umgebung ist, dann flieht der Kimu nicht. Auf rot angemalte Snorfs, würde ein Kimu zulaufen. Tests würden ergeben, dass ein Kimu einen Snorf nicht von einem Loch in der Wand unterscheiden kann

¹⁶³ Eigentlich genauer der Zubeschreibung von intentionalen Zuständen.

¹⁶⁴ „Perhaps my intuitions here are the products of causal theorie of reference and reliabilist theories of knowledge.“ (Pietroski 1992, 277)

(Vgl. Pietroski 1992, 276). Kurzum, Pietroski scheint folgende Annahme, die man als Diskriminationsprinzip bezeichnen könnte, vorauszusetzen:

*Ein Organismus kann keine Überzeugung über X haben, wenn der Organismus nicht in der Lage ist, X von non-X zu unterscheiden.*¹⁶⁵

Wir können dementsprechend auch die Prämisse 5b als Mini-Argument darstellen.

1. Eine intentionale Erklärung ist völlig unplausibel, wenn sie Überzeugungen¹⁶⁶ enthält, die der Organismus, dessen Verhalten erklärt werden soll, nicht haben kann.
 2. Die intentionale Erklärung des Kimu-Verhaltens Millikan Style enthält Überzeugungen über Snorfs.
 3. Ein Organismus kann keine Überzeugung über X haben, wenn der Organismus nicht in der Lage ist, X von non-X zu unterscheiden.
 4. Kimus können Snorfs nicht von nicht-Snorfs unterscheiden.
 5. Kimus haben keine Überzeugungen über Snorfs (aus 3 und 4).
-
6. Die intentionale Erklärung des Kimu-Verhaltens Millikan Style ist völlig unplausibel (aus 1, 2 und 5)

Damit ist meine Rekonstruktion von Pietroskis Argument abgeschlossen. Wir wollen nun das Argument kritisch prüfen.

Kritische Prüfung von Pietroskis Einwand

Für die kritische Prüfung von Pietroskis Einwand möchte ich mich an der Argumentation von Markus Wild orientieren. Ausgehend von Pietroskis „Biologiemärchen“ (Wild 2010a, 354) verteidigt Wild auf einhundert Seiten Millikans Biosemantik gegen Pietroskis Kritik (Vgl. Wild 2010a, 357-457).

Wilds erste Reaktion: mangelnde Differenzierung

Der erste Gegeneinwand von Wild ist, dass Pietroskis Kritik von einer mangelnden Differenzierung profitiert¹⁶⁷. Pietroski scheint nämlich davon auszugehen, dass ein Organismus der Gehaltsträger im Sinne Millikans ausbildet auch automatisch über Überzeugungen des gleichen Gehalts verfügen muss.

„Millikan thus bites the bullet, holding that kimu’s B-tokens are about snorfs. But then kimus climb the hill *because they believe* that the hill is snorf-less [...]“ (Pietroski 1992, 276, Hervorh. Im Original)¹⁶⁸

Diese Annahme ist aber ungerechtfertigt, insbesondere im Rahmen einer Kritik an Millikans Theorie, denn Millikan differenziert durchaus zwischen verschiedenen intentionalen Zeichen. Bei kognitiv weniger entwickelten Lebewesen finden wir nach Millikan nur Repräsentationen eines Typs, den sie als Pushmi-pullyu-Repräsentation (PPR) bezeichnet. Die PPRs sind dadurch charakterisiert, dass sie zugleich eine indikative und imperative Gehaltskomponente haben (Vgl. Abschnitt 2.1.1.2). Demgegenüber versteht Millikan Überzeugungen als rein indikative Zeichen, die durch etliche weitere Besonderheiten charakterisiert sind, wie z.B.

¹⁶⁵ Ich halte diese Annahme für sehr plausibel.

¹⁶⁶ oder Wünsche.

¹⁶⁷ In eine ganz ähnliche Richtung geht der Gegeneinwand von Papineau und Macdonald, wenn sie feststellen, dass die kontra-intuitiven Konsequenzen, davon abhängen mehr in das Kimu-Beispiel hineinzulesen als Pietroskis Beschreibung eigentlich hergibt (Vgl. Macdonald und Papineau 2006, 9).

¹⁶⁸ Bis zu dieser Stelle hatte Pietroski nur von B-tokens und Repräsentationen gesprochen, aber nicht von Überzeugungen.

ihre Negierbarkeit (Vgl. Millikan 1989b, 297)¹⁶⁹. Die B-tokens der Kimus sind nach Millikan eindeutig PPRs und keinesfalls Überzeugungen.

"Eine PPR enthält sowohl imperative als auch indikative Elemente. [...] Aus diesem Grund hat die sprachliche Formulierung des Inhalts der Kimu-Repräsentation die Form "snorffreie Richtung!". Das Ausrufezeichen soll den direktiven Aspekt dieser PPR sichtbar machen." (Wild 2010a, 359)

Es ist aber auch ganz abgesehen von Millikans Unterscheidungen in indikative, imperative Zeichen und PPRs, nicht unbedingt plausibel, geschweige denn zwingend die B-tokens der Kimus als Überzeugungen zu deuten. Es sei denn, wir wollen auch die Ausrichtung der Magnetosomen von Bakterien als Überzeugungen der Bakterien verstehen¹⁷⁰.

Wie profitiert nun Pietroskis Argumentation von dieser mangelnden Differenzierung? Pietroski hatte zunächst ganz korrekt den Gehalt der B-tokens anhand von Millikans Theorie bestimmt. Er deutet dann aber ohne jede weitere Begründung die B-tokens als Überzeugungen (Vgl. dazu die Erläuterung zu den Prämissen 3a, 5a und 5b oben) und kann nur so unsere Intuitionen zu Überzeugungen nutzen um Millikans Gehaltsbestimmung als unakzeptabel erscheinen zu lassen. Das Argument von Pietroski lässt sich nun so wiedergeben:

1. a hat nach Millikans Theorie den Gehalt x
 2. a ist eine Überzeugung
 3. Wenn a nach Millikans Theorie den Gehalt x hat und a eine Überzeugung ist, dann ist Millikans Theorie unakzeptabel
-
4. Millikans Theorie ist unakzeptabel

Der Gegeneinwand von Wild blockiert also die Konklusion, indem er die zweite Prämisse bestreitet. Wenn die B-tokens der Kimus keine Überzeugungen sind, dann spielen unsere Intuitionen zu Überzeugungen auch keine Rolle dafür, welchen Gehalt die B-tokens haben oder auch nicht haben können. Mit anderen Worten: Selbst wenn Pietroskis Intuitionen zu Überzeugungen korrekt sind, ergeben sich daraus keine Konsequenzen für die Art und Weise wie Millikan den Gehalt der B-tokens bestimmt, einfach weil die B-tokens keine Überzeugungen sind.

Pietroskis Einwand kann also nicht überzeugen. Ist der Einwand damit aber völlig entkräftet? Sicherlich nicht vollständig, denn man kann den Einwand modifizieren, so dass an die Stelle von Überzeugungen Wahrnehmungen treten. Abzustreiten, dass Kimus Wahrnehmungen haben, ist sicherlich schwieriger als abzustreiten, dass Kimus Überzeugungen haben. In modifizierter Form würde das Argument lauten:

1. a hat nach Millikans Theorie den Gehalt x
 2. a ist eine Wahrnehmung
 3. Wenn a den Gehalt x hat und a eine Wahrnehmung ist, dann ist Millikans Theorie unakzeptabel
-
4. Millikans Theorie ist unakzeptabel

¹⁶⁹ Millikan hat immer wieder betont, dass Überzeugungen ganz spezifische Zeichen sind (Vgl. z.B. Millikan 1984, 71, 96; Millikan 1989b, 97-101; Millikan 2008, 219-310).

¹⁷⁰ Das Kimu-Beispiel wird von Pietroski selbst als analoger Fall zu Dretskes Bakterien-Beispiel behandelt (Vgl. Pietroski 1992, 279f).

Man könnte allerdings auch gegen diesen Zug wieder einwenden, dass Wahrnehmungen oft als rein indikativisch verstanden werden und die B-tokens der Kimus, wie bereits festgestellt, PPRs sind. D.h. auch bei dieser Modifikation läuft man leicht Gefahr in das Kimu-Beispiel mehr hineinzulesen als eigentlich beschrieben wurde. Auf diese Gefahr und den möglichen Ausweg für Millikan weisen auch Macdonald und Papineau hin:

„What about the counter-intuitive consequences that Pietroski's kimu thought-experiment seems to foist on output-based versions of teleosemantics? Perhaps there is room for output-based teleosemantic theories to argue that these intuitions depend on reading more into Pietroski's scenario than is justified by his description. Pietroski's says that the kimu evolve some state that is triggered by redness and which has the advantage of keeping them away from the snorf. Given this specification, it is hard to stop ourselves thinking of the kimu as having some general-purpose visual *system* which gathers items of visual information which might then be used to inform an open-ended range of behavioral projects directed at different possible ends [...]. However, this extra structure in fact takes us significantly beyond what Pietroski's description actually requires, and it is open to output-based teleosemanticists to argue that their theory is quite able to explain why an organism with all this extra structure would be representing redness rather than snorf-freeness: if the organism's visual states inform a *range* of different behaviors directed at different ends, then the content of any such state needs to be fixed as some condition that assists in achievement of *all* those ends, and this may well come out as redness. On the other hand, if we do stick to a minimal understanding of the kimu, as having only a special-purpose visual sensitivity that brings no advantage except snorf-avoidance, then it's not so clear that there is anything wrong with the output-based reading of their states as representing *snorf-freeness*: after all, if these states never do anything except trigger simple avoidance behaviour, it seems natural enough to read them as representing the danger they are designed to avoid.“ (Macdonald und Papineau 2006, 9)

Somit würde auch das modifizierte Argument scheitern, weil man wieder die zweite Prämisse bestreiten könnte. Diese Verteidigung wäre allerdings etwas schwächer, da sie sich auf Millikans Einteilung in indikative, imperative und Pushmi-Pullyu Repräsentationen stützen muss. Millikans Einteilung ist zwar nicht unplausibel aber auch nicht sakrosankt. Wir wollen deshalb das Bedenken, dass Wahrnehmungen keine PPRs sind, im Folgenden unberücksichtigt lassen und sehen, was Wild dem modifizierten Einwand entgegensetzen hat.

Wilds zweite Reaktion¹⁷¹: Schwächung von Pietroskis Intuitionen

Wir haben gesehen, dass es zumindest fragwürdig ist einige unserer Intuitionen zu Überzeugungen und Wahrnehmungen – so berechtigt diese auch sein mögen – auf das Kimu-Beispiel anzuwenden. Wild versucht nun in einem zweiten Schritt unsere Intuitionen selbst in Frage zu stellen. Die Intuition von Pietroski, dass zwischen Snorfs und B-tokens ein kausaler Zusammenhang bestehen muss, damit die B-tokens von Snorfs handeln können, bezeichnet Wild allgemein als das Nomische Korrelationsprinzip (NKP):

¹⁷¹ Eigentlich ist dies Wilds dritte Reaktion. Seine zweite Reaktion ist zu zeigen, dass eine produzentenorientierte Version der Teleosemantik (Vgl. Dretske 1988, Neander 2006), die Pietroskis Einwand scheinbar zu entgehen scheint, letztlich nicht besser sondern schlechter abschneidet als die Biosemantik Millikans (Vgl. Wild 2010a, 361-406). Da dies nicht unmittelbar zur Entkräftung von Pietroskis Einwand beiträgt, soll dieser Teil von Wilds Argumentation hier unberücksichtigt bleiben. Pietroski überlegt zwar, ob man Millikans Theorie nicht trotz ihrer kontraintuitiven Konsequenzen akzeptieren sollte, weil wir eben keine bessere Theorie haben (Vgl. Pietroskis 1992, 276). Bezüglich dieser Überlegung von Pietroskis könnte Wilds Kritik der produzentenorientierten Teleosemantik dann doch beitragen. Schliesslich tendiert Pietroski aber eher dazu die gesamte biosemantische Idee zu verwerfen als die kontra-intuitiven Konsequenzen zu akzeptieren.

„**NKP**: Wahrnehmungen können eine Eigenschaft *F* nur intentional repräsentieren, wenn die *R*-Vehikel mit Instantiierungen von *F* nomisch korreliert sind.“ (Wild 2010a, 406)

Wild versucht nun zu zeigen, dass weder das Nomische Korrelationsprinzip noch der Ansatz bei der Wahrnehmungserfahrung¹⁷² notwendige Bestandteile einer Theorie der Wahrnehmung sind.

„Die tiefste Ebene des Einwands aus Pietroskis Biologiemärchen kann in diese Frage gefasst werden: Verzichtet die Biosemantik mit NKP nicht auf zu viel? Ich möchte dem widersprechen und die These vertreten, dass eine biosemantische Theorie der Wahrnehmung nicht nur NKP als notwendigen Bestandteil einer Analyse des Wahrnehmungsbegriffs fahren lassen darf, sondern auch den Ansatz bei Wahrnehmungserfahrungen und die Idee einer Analyse des Wahrnehmungsbegriffs.“ (Wild 408)

Die generelle Strategie, die Wild dabei verfolgt, ist es Gegenintuitionen zu erzeugen, indem er fiktive Fälle oder Grenzfälle beschreibt (Vgl. Wild 2010a, 413).

Schauen wir uns also zuerst an, wie Wild das Vertrauen in das Nomische Korrelationsprinzip untergraben will. Wild entwickelt seine Überlegungen hier in Hinblick auf die kausale Theorie der Wahrnehmung, deren Grundlagen Paul Grice (1989) legt hat. Die kausale Theorie der Wahrnehmung erfasst dabei insofern unsere Intuitionen zur Wahrnehmung, wie das Nomische Korrelationsprinzip, als sie sich als eine begriffliche Analyse des vortheoretischen Wahrnehmungsbegriffs versteht (Vgl. Wild 2010a, 408). Da Wild sich nur für die visuelle Wahrnehmung interessiert, spricht er oft nur von einer kausalen Theorie des Sehens (KTS). Wild zufolge gilt nach der kausalen Theorie des Sehens Folgendes:

„*S* sieht *o* als *F* gdw gilt:

- (1) *S* hat eine visuelle Erfahrung *E* von *o* als *F*,
- (2) *o* ist *F*,
- (3) *E* hängt kausal von der Tatsache ab, dass *o* *F* ist.“ (Wild, 409f)

In der dritten Bedingung ist das Nomische Korrelationsprinzip erfasst¹⁷³. Um nun eine Gegenintuition gegen dieses Prinzip zu erzeugen, führt Wild den Fall des blinden Geronimo an:

"Das Hirn des vollkommen blinden Geronimo wird mit einem Computer so verbunden, dass ein Neurochirurg in Geronimo die visuellen Halluzinationen von just jenen Szenarien hervorzurufen vermag, die sich tatsächlich vor seinen blinden Augen abspielen. [...] Der Neurochirurg garantiert die attitudionale kontrafaktische Abhängigkeit der halluzinierten visuellen Wahrnehmung Geronimos von der optischen Szenerie vor dessen Augen. [...] Auch die kausale Abhängigkeit ist gegeben. Geronimo sieht die Szenerie vor seinen Augen in attitudinaler kontrafaktischer Abhängigkeit, nämlich in Abhängigkeit sowohl von den Szenarien vor seinen Augen als auch von der Auffassung der Szenarien, die der Neurochirurg ausbildet." (Wild, 410)

¹⁷² Der Ansatz bei der Wahrnehmungserfahrung spielt in Pietroskis Argumentation keine explizite Rolle, dennoch scheinen Aspekte bewussten Wahrnehmungserfahrung seiner Argumentation intuitive Kraft zu verleihen. Wir werden darauf später noch zurückkommen.

¹⁷³ "Für das Sehen von *o* bzw. *F* oder für das Sehen, dass *o* *F* ist, ist es nicht nur eine notwendige Bedingung, dass *o* bzw. *F* vorhanden sind oder dass *o* *F* ist, sondern ebenso, dass die visuelle Erfahrung *E* von *o*, *F* oder *oF* nicht wäre, wie sie ist, wenn *o* bzw. *F* nicht vorgelegen hätte oder *o* nicht *F* wäre. Wie die Dinge visuell erscheinen hängt (im veridischen Falle) davon ab, wie die Dinge sind. Und diese Abhängigkeit muss kontrafaktisch sein. Insofern respektiert KTS also NKP." (Wild, 410)

Die kausale Theorie des Sehens will Wild zufolge den Fall des blinden Geronimo als genuinen Falls des Sehens ausschliessen, weil es sich in diesem Fall nicht um den „*normalen kausalen Weg* von den Szenerien zum Wahrnehmungserlebnis handelt“ (Wild 2010a, 410). Um den Ausschluss zu bewirken, muss die dritte Bedingung derart modifiziert werden, dass sie besagt, dass *E* auf *natürliche Weise* oder unter *normalen Bedingungen* von *o* oder *F* verursacht wird (Vgl. Wild 2010a, 411). Diese Modifikation wiederum generiert zwei Probleme, erstens ist die Bestimmung dessen, was normal bzw. natürlich ist, von empirischen Befunden abhängig¹⁷⁴ und zweitens werden Fälle ausgeschlossen, die nach unserem Alltagsverständnis genuine Fälle des Sehens sind, wie das Sehen mittels Spiegeln, Brillen, Teleskopen und Mikroskopen (Vgl. Wild 2010a, 411). Dies ist im Wesentlichen Wilds Argumentation gegen das Nomische Korrelationsprinzip und damit letztlich gegen Pietroskis Intuition, dass zwischen Snorfs und B-tokens ein kausaler Zusammenhang bestehen muss, damit die B-tokens von Snorfs handeln können.

Meines Erachtens kann Wilds Gedankengang hier nicht restlos überzeugen, weil er zu wenig zeigt. Der Fall des blinden Geronimo zeigt, dass es Fälle gibt, die nicht eindeutig zu beantworten sind. Sieht der blinde Geronimo die Szenerien vor sich oder sieht er sie nicht? Aus der Existenz solcher Fälle scheint dann zu folgen, dass der Versuch einer trennscharfen Definition von „Sehen“ problematisch ist. Was Wild bestenfalls gezeigt hat, ist also, dass der Versuch den Fall des blinden Geronimo als genuinen Wahrnehmungsfall auszuschliessen in ernsthafte Schwierigkeiten gerät. Dieser Punkt wäre für die Gesamtargumentation von entscheidender Bedeutung, wenn der Fall des blinden Geronimo zugleich ein Fall wäre, in dem zwischen den Szenerien und Geronimo bzw. dessen Wahrnehmungserlebnissen kein kausaler Einfluss bestünde, was aber nicht der Fall ist. Es besteht ja wie Wild selbst schreibt (siehe Zitat oben) eine kausale Verbindung über den Wissenschaftler. Wild zeigt also höchstens ganz allgemein, dass unsere Alltagsintuitionen zur Wahrnehmung zu Schwierigkeiten führen, wenn wir aus ihnen trennscharfe Definitionen ableiten wollen. Dies wiederum deutet daraufhin, dass unsere Intuitionen in diesem Bereich weniger kohärent sind als wir meist annehmen, was uns letztlich auch dazu bewegen könnte uns nicht vollends auf diese Intuitionen zu verlassen, wenn wir Theorien wie Millikans Biosemantik beurteilen. Was Wild nicht gezeigt hat, ist, dass die spezielle Intuition von Pietroski, nämlich dass zwischen Snorfs und B-tokens ein kausaler Zusammenhang bestehen muss, damit die B-tokens von Snorfs handeln können, auf der Pietroskis Argument beruht, problematisch ist. Dazu hätte der Fall des blinden Geronimo so konstruiert werden müssen, dass zwischen Geronimo und den Szenerien kein kausaler Zusammenhang besteht. Erst dann hätten wir eine echte Gegenintuition zu Pietroskis Auffassung erzeugt. Wild bleibt hier also etwas hinter seinen Ankündigungen zurück.

Schauen wir uns nun an, wie Wild gegen den Ansatz der Wahrnehmungstheorien bei der Wahrnehmungserfahrung argumentiert. Bevor wir dies tun, müssen wir aber noch kurz erläutern, inwiefern unsere Intuitionen zum Kimu-Beispiel mit der Wahrnehmungserfahrung verbunden sind. Die grundlegende Intuition zur Wahrnehmungserfahrung, um die es Wild geht, ist die Folgende:

„Es ist sicher nicht falsch zu sagen, ein Subjekt sehe, wenn die Dinge oder die Szenerie vor seinen Augen in ihm direkt entsprechende, passende visuelle Erfahrungen verursachen.“ (Wild 2010a, 411)

¹⁷⁴ Wo diesbezüglich das Problem liegen soll, wird von Wild nicht richtig erläutert. Er meint lediglich, dass dieses Problem durch die „Reinheit der Begriffsanalyse“ (Wild 2010a, 411) erledigt wird, was vermutlich ironisch gemeint ist und darauf hindeutet, dass die Bestimmung dessen, was normal bzw. natürlich ist, im Rahmen der kausalen Theorie des Sehens nicht geleistet werden kann.

Nach dieser Vorstellung gilt, ich sehe rote Blumen, wenn die roten Blumen vor mir, bei mir eine visuelle ROTE-BLUMEN-ERFAHRUNG verursachen. Zudem gehen wir davon aus, dass wir die Erfahrungen verbal beschreiben können (siehe Abbildung 62).

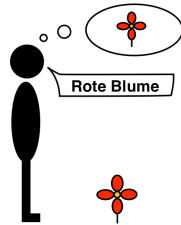


Abbildung 62: Wahrnehmungserfahrung, Beschreibung der Wahrnehmungserfahrung und wahrgenommenes Objekt bei Menschen

Wenn wir diese Vorstellung auf das Kimu-Beispiel übertragen, dann erwarten wir, dass der Kimu Jack in der Gegenwart von roten Blumen ebenfalls eine visuelle ROTE-BLUMEN-ERFAHRUNG hat und dass die Beschreibung dieser Erfahrung eben „Rote Blumen, so und so beschaffen ...“ und keineswegs „Snorffreie Zone, so und so beschaffen ...“ sein muss. Wenn wir zudem annehmen, dass der Gehalt der visuellen Erfahrung dem Gehalt der internen Repräsentation (b-tokens) entsprechen muss oder man gleich die visuelle Erfahrung mit den B-tokens identifiziert, dann folgt daraus, dass Millikans Gehaltsbestimmung falsch ist.

Dieser Argumentationsansatz findet sich nicht explizit bei Pietroski. Es ist aber anzunehmen, dass unsere Intuitionen bei der Beurteilung des Kimu-Beispiels auch von den eben genannten Vorstellungen beeinflusst werden. So hatte Hall zwei Jahre vor Pietroski anhand von Beispielen, die Pietroskis Kimu-Beispiel ganz ähnlich sind, versucht zu zeigen, dass bewusste Wahrnehmungserfahrungen im Konflikt mit Millikans Gehaltszuschreibungen stehen (Vgl. Hall 1990). Hall bringt folgendes Beispiel:

„Consider a kind of worm that has evolved in an environment where all and only its predators cast moving shadows. In worm world, „predator“ and „casts moving shadows“ are extensionally equivalent. Isn't it natural to suppose the worm detects and represents moving shadow even though there's no biological law that the avoidance of moving shadow casters are selected for? Representing and avoiding casters of moving shadows is how the worm achieves the avoidance of predators.“ (Hall 1990, 197)

Für Hall scheint es völlig klar zu sein, dass der Wurm bewegte Schatten und nicht Prädatoren repräsentiert, dass also bestimmte interne Repräsentationen des Wurms den Gehalt „bewegter Schatten“ und nicht „Prädatoren“ haben. Der Hauptgrund für diese Überzeugung scheint folgende Überlegung zu sein: Wenn der Wurm zu bewussten Wahrnehmungen fähig wäre, dann würde er seine Wahrnehmungserfahrung als ein Wahrnehmen eines Schattens und nicht eines Prädatoren beschreiben. Dieser Überlegung wiederum liegt die Annahme zugrunde, dass der Gehalt einer internen Repräsentation eng mit dem verbunden ist, was der Organismus, der die interne Repräsentation hat, bewusst wahrnimmt.

„Our original notion of representational content was connected, at least in our human case and that of higher animals, with what we are consciously aware of. What we are led to on the proposed interpretation of the biological function account is a notion of content that may have no connection (except local coextensiveness) with conscious content. The biological functionalist would describe the content of the internal state of the worm which causes it to crawl down its hole when a moving shadow falls across it (= when a predator is nearby) as: "predator nearby". But the worm, if it had conscious states and could talk, might say: "What I was consciously aware of, what I perceived,

was a moving shadow. And I believe that its best to crawl down my hole when I perceive a moving shadow.”“ (Hall 1990, 198)

Diese Argumentation entspricht derjenigen, die wir zuvor auch für die Kimus vortragen haben. Die Bedeutung des Ansatzes bei der bewussten Wahrnehmungserfahrung dürfte damit hinreichend klar sein.

Wir können uns somit wieder der Argumentation von Wild zuwenden. Wild geht es bei seiner Argumentation um die grundlegende Frage, welche Bedeutung der bewussten Erfahrung in Wahrnehmungstheorien überhaupt zukommen sollte. Seiner Meinung nach spielt die bewusste Erfahrung keine entscheidende Rolle. Dementsprechend versucht er anhand von zwei weiteren Beispielen zu zeigen, dass das Haben der richtigen bewussten visuellen Erfahrungen nicht entscheidend dafür ist, ob wir jemanden zugestehen *dass* und *was* er sieht.

Wild erstes Beispiel ist der aus den Qualia-Debatten hinlänglich bekannte Zombie. Der Zombie hat per definitionem keine bewussten visuellen Erfahrungen, sein Verhalten ist aber von dem eines Menschen mit solchen Erfahrungen nicht zu unterscheiden¹⁷⁵. Vor einer roten Blume stehend und danach gefragt, was er sähe, würde der Zombie antworten: „Eine rote Blume“ (siehe Abbildung 63).

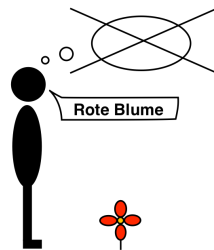


Abbildung 63: Lautäußerung und wahrgenommenes Objekt ohne Wahrnehmungserfahrung beim Zombie

Der Zombie ist für Wild zudem nur die Zuspitzung eines ganz realen Falls, nämlich der Blindsightigkeit¹⁷⁶. Nach Wild kommen wir nicht umhin den Zombies und den Blindsightigen zuzugestehen, dass sie etwas (die rote Blume) sehen, weil wir selbst keine anderen Beschreibungen liefern würden, kein anderes (verbales) Verhalten zeigen würden.

„Der Grund für diese Zuschreibungen besteht darin, dass wir nicht umhin können, die Beschreibungen, die die angeführten Figuren von den visuellen Szenen geben, als Beschreibungen von gesehenen Szenerien anzuerkennen, weil auch wir Normalsichtigen solche Beschreibungen von denselben visuellen Szenen geben würden.“ (Wild 2010a, 413)

Ähnliches soll auch für Wilds zweites Beispiel, die Olfaktoren, gelten. Olfaktoren sind fiktive Lebewesen, die an Stelle von visuellen Erfahrungen olfaktorische Erfahrungen haben. D.h.:

¹⁷⁵ “Zombies: beings physically identical to humans but lacking phenomenal consciousness. If we can imagine them without any contradiction, they are conceivable and thus possible. David Chalmers recently resurrected the zombie idea as an argument against physicalism intended to show that the identity of mental states with brain states cannot be necessary because there is no a priori entailment from physical to phenomenal facts. Critics question the move from conceivability to possibility or address the special nature of phenomenal concepts.” (Rakova 2006, 194).

¹⁷⁶ “Blindsight: the condition (cortical blindness resulting from the destruction of primary visual cortex) in which subjects have no experience of an object presented to their blind field (scotoma) and are incapable of identifying it. [...] However, they retain the ability to locate ‘unseen’ objects and discriminate between them if prompted to guess from a small number of alternatives. (Rakova 2006, 17).

„Wenn Olfaktoren einen roten Apfel zu Gesicht bekommen, haben sie olfaktorische Erfahrungen eines Typs, die ein normales Exemplar unserer Art hätte, wenn es schwefelhaltigen Wasserdampf riechen würde. Olfaktoren urteilen aufgrund dieser olfaktorischen Erfahrung, dass der Apfel rot ist, wie wir aufgrund unserer visuellen Erfahrung urteilen, dass der Apfel rot ist.“ (Wild 414)

Ansonsten Verhalten sich Olfaktoren aber ganz genauso wie sehende Menschen (siehe Abbildung 64).

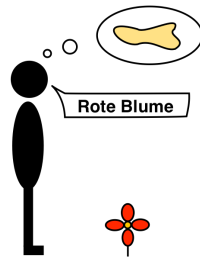


Abbildung 64: Wahrnehmungserfahrung, Beschreibung der Wahrnehmungserfahrung und wahrgenommenes Objekt bei Olfaktoren

Olfaktoren, Blindsichtige und Zombies zeigen Wild zufolge, dass die bewusste visuelle Erfahrung nicht der entscheidende Faktor für unsere Zuschreibung ist, dass ein Subjekt etwas sieht. Entscheidend ist das Verhalten und die Aussage des Subjektes.

Inwiefern hilft dies nun aber um Pietroskis bzw. Halls Argumentation zu entkräften? Auf den ersten Blick gar nicht. Müsste der Kimu nach Millikans Theorie nicht angesichts einer roten Blume sagen, er sähe eine snorffreie Gegend vor sich? Dass er dies tatsächlich sieht, würden wir ihm aber nun nach Wilds Auffassung absprechen müssen, weil wir in etlichen Situationen (siehe Abbildung 65) zu ganz anderen Beschreibungen gelangen würden als der Kimu¹⁷⁷. Wenn sich etwas Rotes und Snorfs im Sichtfeld befinden, dann behauptet der Kimu im Gegensatz zu uns, dass die Gegend snorffrei sei. Befinden sich hingegen keine roten Objekte und keine Snorfs im Sichtfeld, dann behauptet der Kimu, Snorfs seien anwesend, während wir dies korrekterweise verneinen.

Objekte im Sichtfeld	Snorffreie Gegend?	
	Wir	Kimu
rot / Snorf	nein	ja
rot / -	ja	ja
- / Snorf	nein	nein
- / -	ja	nein

Abbildung 65: Antworten, die wir bzw. die Kimus auf die Frage „Ist die Gegend snorffrei?“ in bestimmten Situationen geben würden

Dies würde aber dafür und nicht dagegen sprechen, dass etwas mit Millikans Gehaltszuschreibung nicht stimmt, genau wie es Pietroski und Hall behaupten.

Wir müssen hier aber wieder aufpassen, dass wir nicht viel mehr in das Kimu-Beispiel hineinlesen als uns durch Pietroskis Beschreibung gegeben wurde. Man könnte also gegen eine Nutzung von Halls Argumentation im Kimu-Fall einwenden, dass es absolut

¹⁷⁷ In ähnlicher Weise hatte schon Pietroski argumentiert, dass die Kimus einen psychologischen Test, der deren Fähigkeit Snorfs zu erkennen, nicht bestehen würden (Pietroskis 1992, 276).

ungerechtfertigt ist die Kimus aufgrund von Pietroskis Beschreibung als bewusstseinsfähige Lebewesen zu behandeln¹⁷⁸.

Darüber hinaus ist die Annahme problematisch, dass der Gehalt der visuellen Erfahrung dem Gehalt der internen Repräsentation (B-tokens) entsprechen muss bzw. dass die visuelle Erfahrung mit den B-tokens identifiziert werden kann. Bewusste Erfahrungen, das, was Husserls als Erlebnisse bezeichnet hat¹⁷⁹, sind keinesfalls mit Millikans intentional icons gleichzusetzen, noch stehen die intentional icons mit bewussten Erfahrungen in einer einfachen, direkten Beziehung. Die Gefahr einer phänomenologischen Deutung der internen Repräsentationen der Kimus (B-tokens) bemerkt auch Millikan und weist diese klar zurück.

„Now first, we should note that the injection of phenomenology here is perversely distracting. The question is not whether a red qualia, should there exist such things, could mean no snorfs this direction rather than red. [...] Pietroski's question should be whether any inner representation that merely directs the kimu towards the sunlight could represent for it the snorf free direction.“ (Millikan 2000, 236)

D.h. vor allem, dass nicht jeder internen Repräsentation in Sinne von Millikans intentional icon eine bewusste Erfahrung entsprechen muss, die denselben Gehalt hat wie die interne Repräsentation¹⁸⁰. Auf dieser Annahme beruht aber die Vorstellung der Kimu müsste behaupten, er sähe eine snorffreie Gegend vor sich.

Dass diese Annahme wenig Sinn ergibt, wird besonders deutlich, wenn wir uns daran erinnern, dass Millikans Theorie den Homunkulus-Vorwurf beseitigen soll (Vgl. Millikan 1984, 89). Es ist mit Millikans Theorie absolut vereinbar, dass ein Organismus gleichzeitig viele verschiedene interne Repräsentationen haben kann. Die Annahme, dass diesen jeweils eine bewusste Erfahrung mit demselben Gehalt entsprechen müsse, würde geradewegs in den Homunkulismus führen.

Richtig ist allerdings, dass ultimativ erklärt werden muss, wie bei Lebewesen, die bewusste Erfahrungen haben können, der Gehalt dieser Erfahrung mit dem Gehalt der internen Repräsentationen zusammenhängt. Millikan erklärt dies allerdings nirgends, noch strebt sie eine solche Erklärung überhaupt an.

"What consciousness consists in from a naturalist's point of view, if it consists in anything from a naturalist's point of view, I have nothing to say about." (Millikan 1984, 48)

Dies hängt auch mit Millikans Überzeugung zusammen, dass Intentionalität völlig unabhängig vom Bewusstsein verstanden werden kann:

„For, I will argue, the problem of understanding intentionality can and should be divorced from the problem of understanding consciousness. This is done by abandoning the traditional epistemic view of consciousness by giving up the rationalist view of meaning and intentionality.“ (Millikan 1984, 12)

¹⁷⁸ Hall ist sich dessen bewusst: „Perhaps worms don't have conscious mental contents, in which case the divergence I am pointing to won't arise in their case.“ (Hall 1990, 198).

¹⁷⁹ Erlebnisse sind der zentrale Untersuchungsgegenstand von Husserl Phänomenologie. Unter Erlebnissen versteht Husserl „diejenigen Vorkommnisse, zu denen ich einen epistemisch privilegierten Zugang habe: Ein Ereignis ist genau dann eines meiner Erlebnisse, wenn es Eigenschaften hat, von denen gilt: ich kann introspektiv mit Evidenz urteilen (ich kann „adäquat wahrnehmen“), daß sie ihm zukommen.“ (Künne 1986, 174).

¹⁸⁰ Zudem sollte man bewusste Erfahrungen nicht ohne Weiteres mit Qualiaeindrücken gleichsetzen. Es gibt viele Philosophen (z.B. Peter Hacker und Hans-Johann Glock), die bestreiten würden, dass so etwas wie Quale gibt, aber die Existenz von bewussten Erfahrungen für selbstverständlich halten.

Hintergrund dieser Überzeugung ist die bereits in Abschnitt 1.4.3 angesprochene Kritik von Millikan am sogenannten Bedeutungsrationalismus, deren Prüfung nicht Gegenstand dieser Untersuchung ist.

Den grundlegenden Einwand von Hall, dass die Gehalte von bewussten Erfahrungen nicht von Millikans Theorie erfasst werden, antizipiert Millikan bereits 1984 in ihrem Hauptwerk *Language, Thought and other Biological Categories*.

„my point is that there is no guarantee that this conscious mental content will have anything (except local coextensiveness) to do with the representational content now being stipulated by the biological function account.“ (Hall 1990, 198)

„Offer whatever account you like even of the relations between inner representations, conceived as natural-world objects (say, brain states) and the world outside. Whatever these interesting relations are, they are merely external relations. [...] The basic and interesting kind of intentionality is that which we experience in consciously knowing, believing, intending, etc., and this cannot possibly be the same as the external ‚intentionality‘ of which you, the naturalist, propose to offer an analysis. The basic and interesting kind of intentionality is somehow ‚in‘ consciousness, not outside it or externally between it and the world. What exactly is the force here of the notion that we experience the intentionality of our thoughts or consciousness – that intentionality is somehow ‚in‘ consciousness? [...] The view of my adversary, then, is the view that consciousness is transparently and infallibly epistemic. Consciousness is or essentially involves an infallible kind of direct knowing – a knowing that is guaranteed as such from within consciousness itself. [...] Intentionality is a ‚given‘. Hence this intentionality cannot consist of mere matter-of-fact, even of naturally necessary, external relations to the world.“ (Millikan 1984, 90f.)

Was durch dieses längere Zitat deutlich werden soll, ist lediglich, dass Millikan hinter Halls Einwand den Bedeutungsrationalismus vermuten wird, dessen Grundannahmen sie vehement ablehnt (Vgl. Abschnitt 1.4.3). Es soll und muss in dieser Arbeit offen bleiben, ob Millikans Kritik am Bedeutungsrationalismus berechtigt ist oder nicht. Ebenso muss offen bleiben, ob es nicht möglich ist einen theoretischen Zusammenhang zwischen dem Gehalt verschiedener interner Repräsentationen und dem Gehalt einer bewussten Erfahrung herzustellen. Diese Fragen können im Rahmen dieser Arbeit offen bleiben, weil der Anspruch, den Millikan mit der Biosemantik verbindet, deutlich über das hinausgeht, was wir von der Biosemantik erwarten (Vgl. Abschnitt 1.4.3) und offenen Fragen sich erst in Bezug auf Millikans Anspruch ergeben.

FAZIT:

Pietroskis Einwand kann nicht überzeugen. Insgesamt basiert der Einwand darauf in das Kimu-Beispiel mehr hineinzulesen als es tatsächlich beinhaltet. Weder die Annahme, dass die Kimus Überzeugungen haben, noch die Annahme, dass sie bewusste Wahrnehmungserfahrungen haben, ist gerechtfertigt. Die Intuitionen, die Millikans Gehaltszuschreibung bei den Kimus als absurd erscheinen lassen, setzen aber genau diese Annahmen als wahr voraus.

3.2.3 Keil: Kritik des Naturalismus & Biosemantik: ein degeneriertes Forschungsprogramm

In diesem Abschnitt geht es darum, eine Kritik von Geert Keil an Millikans Theorie des Gehalts zu rekonstruieren und anschließend zu diskutieren (Keil 1993a und 1993b). Keil zufolge scheitert Millikans Biosemantik an ihren naturalistischen Ansprüchen. D.h. Keil glaubt nachweisen zu können, dass die Explikation der grundlegenden Begriffe der Biosemantik intentionales Vokabular enthält oder präsupponiert (Vgl. Abschnitt 1.2.3). Keils Kritik steht somit in einem indirekten Zusammenhang mit dem Problem der Missrepräsentation, denn dieses Problem wird erst innerhalb der naturalistischen Theorien des Gehalts akut. Lässt sich der naturalistische Anspruch nicht aufrecht erhalten, dann stellt sich das Problem der Missrepräsentation auch nicht mehr, und vorgebliche Lösungen werden wertlos. Keil hat seine Kritik sowohl in seiner Dissertation (Keil 1993a) als auch in einem wenig später erschienenen Text (Keil 1993b) vortragen. In der Dissertation behauptet Keil eine Widerlegung der Biosemantik liefern zu können, die Argumentation sei aber auf das Buch verstreut, so dass er die Auseinandersetzung mit Millikan und Papineau bei Gelegenheit nachholen wird (Vgl. 1993a, 98).

„Ich bin überzeugt, die wesentlichen Teilargumente für eine gründliche Widerlegung des Teleofunktionalismus beisammen zu haben, die in dem Nachweis bestünde, dass in dem so sorgfältig aufgebauten vermeintlich evolutionstheoretischen Begriffsapparat doch wieder unanalyisierte intentionalistische Elemente enthalten sind.“ (Keil 1993a, 98)

Die Argumentation in Keil (1993a) und Keil (1993b) basiert auf derselben Strategie. Die Argumentationen unterscheiden sich aber im Detail, daher werden wir uns beide Argumentationen getrennt ansehen.

Zuvor ist es aber hilfreich, sich Keils allgemeine Strategie gegen naturalistische Ansätze klar zu machen. Diese Strategie orientiert sich an folgendem Muster:

Als erstes unterstreicht Keil noch einmal den Anspruch naturalistischer Theorien. Dieser besteht in Übereinstimmung vieler Naturalisten darin, „die fragliche Relation des Bedeutens auf eine solche Relation zurückzuführen, deren Erläuterung weder intentionale noch semantische Begriffe voraussetzt.“ (Keil 1993b, 87) Keil argumentiert anschließend nicht dafür, dass dieser Anspruch prinzipiell nicht einlösbar ist. Sondern er zeigt jeder naturalistischen Theorie einzeln, dass sie ihren eigenen Ansprüchen nicht gerecht wird, indem er der gerade in Frage stehenden Theorie intentionales Vokabular im Explanans nachweist. Keils Motto lautet also: „Liebe Naturalisten, gebt mir nur all eure angeblich naturalistischen Theorien her. Ich weis euch bezüglich jeder Theorie einzeln nach, dass ihr in euren Voraussetzungen zumindest implizit von intentionalen Vokabular Gebrauch gemacht habt.“ Soviel zu Keils Gesamtstrategie.

Rekonstruktion von Keils ersten Argument

Keil erstes Argument lässt sich meines Erachtens so rekonstruieren:

1. Eine Theorie ist naturalistisch ($N(x)$), wenn ihre grundlegenden Begriffe nicht intentionalistisch sind, noch intentionalistisches Vokabular in irgendeiner Form voraussetzen ($\neg I(x)$). [wenn $N(x)$, dann $\neg I(x)$]

2. Millikans Definition der reproduktiv etablierten Familie erster Ordnung enthält den Begriff der Gleichheit.
 3. Millikans Theorie setzt mit dem Begriff der Gleichheit intentionalistisches Vokabular voraus. [I(a)]
-
4. Millikans Theorie ist keine naturalistische Theorie [$\neg N(a)$]

Nachdem wir die Struktur des Argumentes rekonstruiert haben, wollen wir nun die einzelnen Prämissen genauer anschauen.

Die *erste Prämisse* gibt lediglich Keils Verständnis von Naturalismus wieder. Dieses Verständnis habe ich auch für vorliegende Arbeit vorausgesetzt (Vgl. Abschnitt 1.2.3).

Kommen wir zur *zweiten Prämisse*. Millikans Definition der reproduktiv etablierten Familie erster Ordnung enthält in der Tat den Begriff der Gleichheit. Es ist dort die Rede von „gleichen Merkmalen“ (same character) und „gleichen Modellen“ (same model).

„Any set of entities having the same or similar reproductively established characters derived by repetitive reproductions from the same character of the same model or models form a *first-order reproductive established family*.“ (Millikan 1984, 23)

Keil selbst bezieht sich nicht explizit auf die obenstehende Definition. Seine Millikan Referenz (siehe Zitat unten) macht es aber plausibel, dass Keil diese Stelle im Sinn hat.

Die *dritte Prämisse* besagt, dass der Begriff der Gleichheit intentionalistisches Vokabular präsupponiert. Dies ist der Fall, so Keil, weil die Bestimmung von Merkmalen, Zuständen, Prozessen als „gleich“ eine mentalen Abgleich erfordern und diesen Abgleich nicht die Evolution vornehmen kann.

„Der Begriff der Gleichheit ist einer derjenigen Begriffe, die Millikan (1984) in ihrem grossangelegten Versuch der evolutionstheoretischen Naturalisierung von Intentionalität als unproblematisch voraussetzt (vgl. z.B. 23). Wie wir nun sehen, ist auch an dieser Stelle Intentionalität in subtiler Weise involviert. Die Qualifizierung zweier Zustände, Prozesse als „gleich“ erfordert ihren mentalen Abgleich. Weder der Thermostat noch die Evolution können Gleichheiten wahrnehmen.“ (Keil 1993a, 322)

Diskussion von Keils erstem Argument

Das Argument wurde so rekonstruiert, dass der Schluss aus den Prämissen folgt. Eventuell müsste man noch die Prämisse ergänzen, dass der Begriff der reproduktiv etablierten Familie erster Ordnung ein grundlegender Begriff von Millikans Theorie ist, aber dies sollte aufgrund der Darstellung von Millikans Theorie im zweiten Teils dieser Arbeit klar sein. Es bleibt also kritisch zu prüfen, ob auch die Prämissen wahr sind.

Die *ersten beiden Prämissen* scheinen mir unbedenklich. Die erste Prämisse enthält das Naturalismusverständnis, das wir auch voraussetzten, und die zweite Prämisse ist durch Millikans Definition gut belegt.

Die *dritte Prämisse* ist hingegen problematisch. Ich stimme Keil durchaus zu, dass die Verwendung des Begriffs der Gleichheit in gewisser Weise die Fähigkeit des Abgleichens voraussetzt, aber in diesem Fall ist dies doch die Fähigkeit desjenigen, der die Theorie aufstellt, nicht die Fähigkeit des Systems, dessen intentionale Eigenschaften erklärt werden sollen. In der zitierten Stelle schreibt Millikan nicht einmal implizit die Fähigkeit des Abgleichens einem System oder Prozess zu, auch nicht der Evolution, wie Keil zu meinen scheint (Vgl. Keil 1993, 322). In diesem Sinne erfordert die Verwendung jeglicher Begriffe kognitive Fähigkeiten, so dass nicht zu sehen ist, wie die Vertreter der naturalistischer

Theorien dieser Kritik entkommen könnten¹⁸¹. Um so schlimmer für die Vertreter der naturalistischer Theorien, mag Keil sagen. Aber dass man einem Einwand nicht entgeht ist eine Sache, eine andere Sache ist es, ob der Einwand berechtigt ist. Besteht hier aber Anlass für einen Zirkularitätsvorwurf? Nein, denn eine zirkuläre Erklärung liegt nur dann vor, wenn im Explanans auf intentionale Eigenschaften Bezug genommen wird, nicht aber, wenn man seine kognitiven Fähigkeiten einsetzt, um seine kognitiven Fähigkeiten zu erklären. Ist es nicht einfach schlicht unvermeidlich seine eigenen kognitiven Fähigkeiten zu benutzen, um eine Erklärung seiner eigenen kognitiven Fähigkeiten zu erlangen? Wenn dem so ist, dann kann Millikans Vorgehen nicht dazu führen, dass Millikan den naturalistischen Anspruch ihrer Theorie aufgeben müsste.

Vielleicht will Keil aber auch sagen, dass der Begriff der Gleichheit per se ein intentionaler Begriff ist im Gegensatz zu anderen Begriffen, wie „elektrische Ladung“ oder „Primzahl“. Diesbezüglich muss man aber zunächst klären von welcher Art von Gleichheit hier die Rede ist, von einer numerischen oder von einer qualitativen. Aus dem Millikanzitat geht klar hervor, dass nur der Begriff der qualitativen Gleichheit gemeint sein kann. Keil könnte nun behaupten, es gehe ihm nicht um den Einsatz kognitiver Fähigkeiten zur Erkenntnis der Gleichheit zweier Dinge, sondern darum, ob es Gleichheit oder Ähnlichkeit überhaupt geben kann, ohne dass ein Wesen diese feststellt. Aber falls es Gleichheiten oder Ähnlichkeiten ohne Feststeller gibt, kann man dann nicht (à la Goodman) sagen: Alles ist mit allem irgendwie ähnlich. Welches ist nun die relevante Gleichheit oder Ähnlichkeit, um die es gehen soll? Die muss doch jemand auswählen. Also, so Keil, steckt bereits im Begriff der First-Order Reproductively Established Family Intentionalität (nämlich via den Begriff der Gleichheit). Auch hier ist Keil zunächst zuzustimmen. Zwei Dinge ähneln bzw. gleichen sich immer nur in Hinsicht auf bestimmte Eigenschaften. Aber Hinsicht hat hier nichts damit zu tun, dass jemand auswählend hinsehen muss. Dass zwei Dinge in gewisser Hinsicht gleich sind, heißt einfach nur, dass sie bestimmte Eigenschaften teilen. Und diese teilen sie ganz unabhängig von irgendeinem Beobachter. Natürlich muss bei der Rede von der Gleichheit mehrerer Dinge angegeben werden, in welcher Hinsicht sie sich gleichen. Aber diese Aufgabe übernimmt wiederum derjenige, der die Theorie aufstellt. Der Explanator bestimmt die für seine Erklärung relevanten Eigenschaften. Also läuft auch dieser Versuch eines Einwandes lediglich darauf hinaus, dass der Explanator bei der Formulierung des Explanans seine kognitiven Fähigkeiten zum Einsatz bringt. Aber dies ist, wie wir bereits festgestellt haben, einfach unvermeidlich und daher kein Anlass zur Kritik.

Mit dem vorgestellten Argument gelingt es Keil also nicht seine Strategie umzusetzen und Millikan nachzuweisen, dass einer ihrer Grundbegriffe intentionalistisches Vokabular voraussetzt. Schauen wir uns nun Keils zweites Argument an.

Rekonstruktion von Keils zweitem Argument

Keil verfährt auch hier nach seiner generellen Strategie, deshalb lässt sich sein zweites Argument ganz analog zu seinem ersten Argument rekonstruieren:

1. Eine Theorie ist naturalistisch ($N(x)$), wenn ihre grundlegenden Begriffe nicht intentionalistisch sind, noch intentionalistisches Vokabular in irgendeiner Form voraussetzen ($\neg I(x)$). [wenn $N(x)$, dann $\neg I(x)$]
2. Millikans Theorie setzt mit dem Begriff der Eigenfunktion (proper function) intentionalistisches Vokabular voraus. [$I(a)$]
3. Millikans Theorie ist keine naturalistische Theorie [$\neg N(a)$]

¹⁸¹ In ganz ähnlicher Weise argumentiert auch Walter (1997, 189).

Nachdem wir die Struktur des Argumentes rekonstruiert haben, wollen wir nun wieder die einzelnen Prämissen genauer anschauen.

Während für die *erste Prämisse* wieder das bereits Gesagte gilt, bedarf die *zweite Prämisse* einiger Erläuterung. Zunächst interpretiert Keil die Definition der Eigenfunktion (Vgl. Abschnitt 2.1.1.3.3) als eine naturalistische Explikation des Teleologischen unter Rückgriff auf die biologische Evolution (Vgl. Keil 1993b, 90). Die Frage, ob diese Explikation gelingt, setzt Keil dann mit folgender Frage gleich:

„Inwiefern ist das „um zu“ einer teleologischen Erklärung in das „weil“ einer kausalen Erklärung überführt worden?“ (Keil 1993b, 90)

Keil scheint dabei Folgendes implizit anzunehmen: Wenn sich der teleologische Aspekt (das „um zu“), der mit dem Begriff der Eigenfunktion verbunden ist¹⁸², nicht rein kausal beschreiben lässt, dann setzt der Begriff der Eigenfunktion implizit intentionalistisches Vokabular voraus. Indem Keil also nachweist, dass die kausale Beschreibung misslingt, gedenkt er zu zeigen, dass der Begriff der Eigenfunktion intentionalistisches Vokabular voraussetzt (Prämisse 3). Keil argumentiert nun wie folgt für das Misslingen der kausalen Beschreibung.

Ausgangspunkt seiner Argumentation ist ein Zitat von Papineau:

"when we talk of some characteristic C being present *in order to* produce E, we should understand ourselves to be claim that C is now present because of some past selection process that favoured C because it produced E" (Papineau 1987, 65)

Keil behauptet nun, dass das „because“ im obigen Zitat erschlichen ist (Vgl. Keil 1993b, 90). Hierzu beruft er sich zunächst auf die Kritiken von Nagel und Achinstein an Wrights Formel „X is there *because* it does Z“. Die Kritik von Nagel und Achinstein läuft darauf hinaus, dass ein zukünftiger Effekt keine Ursache für das Vorhandensein eines Funktionsträgers sein kann, einfach weil die Ursache zeitlich vor der Wirkung liegen muss. Obwohl sich nun Millikans und Papineaus Funktionsdefinitionen nicht ausdrücklich auf zukünftige Effekte berufen, glaubt Keil, dass sie dennoch dazu gezwungen sind (Vgl. Keil 1993b, 90). Wenn sie dies nicht täten, dann würde, so Keil, das Prinzip der Natürlichen Selektion, auf das sich beide Definitionen in zentraler Weise stützen, zur Tautologie vom „the survival of the survivor“ (Vgl. Keil 1993b, 91) und dies wäre für Millikan und Papineau unakzeptabel. Mit anderen Worten: Auch Papineau und Millikan müssen ihren Funktionsbegriff so verstehen, dass er sich auf zukünftige positive Effekte bezieht. Genau dies führt aber dazu, dass Papineau und Millikan von Nagel und Achinsteins Kritik betroffen sind. Dementsprechend schlussfolgert Keil:

„Die angestrebte kausalistische Reduktion des „um zu“ scheitert am *dispositionalen* und damit *antizipatorischen Moment* teleologischer oder funktionaler Erklärungen, welches in einer Kausalerklärung nicht einholbar ist.“ (Keil 1993b, 91)

Diskussion von Keils zweiten Argument

Das Argument wurde so rekonstruiert, dass der Schluss aus den Prämissen folgt. Es bleibt also kritisch zu prüfen, ob auch die Prämissen wahr sind.

Die *erste Prämisse* ist wieder unbedenklich.

¹⁸² Das Herz ist dazu da, *um* Blut zu pumpen = Herzen haben die Eigenfunktion Blut zu pumpen.

Die *zweite Prämisse* hingegen scheint fragwürdig. Zunächst ist die Annahme von der Keil auszugehen scheint, nämlich dass der Begriff der Eigenfunktion implizit intentionalistisches Vokabular voraussetzt, wenn sich der teleologische Aspekt, der mit dem Begriff der Eigenfunktion verbunden ist, nicht rein kausal beschreiben lässt, zwar nicht unplausibel aber doch erläuterungsbedürftig. Ich werde hier aber davon ausgehen, dass diese Annahme korrekt ist, weil ich glaube, dass Keils Argumentation ohnehin nicht überzeugend ist.

Weiterhin ist es etwas merkwürdig, dass Keil sich nicht an Millikans Definition der Eigenfunktion orientiert (Ist sie ihm zu lang?), obwohl er sich zuvor im Text fast ausschließlich auf ihre Theorie bezieht, sondern an Papineaus Definition.

Auf den ersten Blick kann zudem der Eindruck entstehen, dass Keil Millikan fälschlicherweise einen dispositionalen Funktionsbegriff unterstellt und seine Argumentation schon deshalb verfehlt ist (Vgl. Walter 1997, 189). Hierzu ist Folgendes zu sagen: Millikan vertritt zwar in der Tat keinen dispositionalen Funktionsbegriff, aber sie tut dies nicht, weil sie glaubt, der Begriff der Disposition präsupponiere intentionalistisches Vokabular, sondern weil ein dispositionaler Funktionsbegriff das Verfehlen einer Funktion und damit eine Lösung des Problem der Missrepräsentation unmöglich macht (Vgl. Abschnitt 3.2.9). Unter einer Disposition versteht man allgemein: „a tendency of an object or system to act or react in characteristic ways in certain situations.“ (Sapire 1999, 238) Der kausale Status von Dispositionen ist zwar umstritten (Vgl. Rakova 2006, 48-49), aber es ist nicht erkennbar, inwiefern der Begriff der Disposition zwingend intentionalistisches Vokabular präsupponiert, zumindest dann nicht, wenn man eine Auffassung von Dispositionen vertritt, wie Prior, Pargetter und Jackson (1982). Mit anderen Worten: Keils Einwand basiert nicht darauf Millikan einen dispositionalen Funktionsbegriff zu unterstellen. Ausschlaggebend für Keils Einwand ist vielmehr seine Überlegung, nach der sich Papineau und Millikan in ihren Definitionen von Funktionen nicht nur auf vergangene Prozesse beziehen können. Genau dies tun sie nämlich (Vgl. das Papineau-Zitat oben und Millikan Definition der Eigenfunktion auf S. 59). Nur dann kann Keil die vollkommen berechtigte Kritik von Nagel und Achinstein an Wright auch auf Papineau und Millikan übertragen. Die entscheidende Frage ist also, ob die Beschränkung auf vergangene Selektionsprozesse tatsächlich den Tautologieeinwand vom „survival of the survivor“ nach sich zieht, wie Keil behauptet, so dass Millikan und Papineau gezwungen wären die Beschränkung aufzuheben, was sie wiederum der Kritik von Nagel und Achinstein anheim fallen lässt.¹⁸³ Hierbei unterstelle ich, wie Keil, dass der Tautologieeinwand vom „survival of the survivor“ unbedingt zu vermeiden ist. Dieser Frage wollen wir nun nachgehen.

Zunächst müssen wir verstehen, worin eigentlich der Tautologieeinwand vom „survival of the survivor“ besteht. Der Einwand bezieht sich auf Darwins Bestimmung der Natürlichen Selektion als dem Überleben des Fittesten.

„This principle of preservation or the survival of the fittest, I have called Natural Selection.“ (Darwin, *The Origin of Species* 6th Edition, Chapter 4, Summary, zitiert nach Eigen 1975, 72)

Solange – so der Einwand – der am besten Angepasste (the fittest) nur dadurch bestimmt ist, dass er überlebt hat, besteht Darwins Prinzip aus einer Tautologie, dem „survival of the survivor“ (Vgl. Eigen 1975, 74).

Es gibt tatsächlich Selektionsprozesse, in denen der Überlebende allein dadurch charakterisiert ist, dass er überlebt hat. Man spricht hier von einer neutralen oder nicht-darwinischen Selektion beziehungsweise von genetischen Drift (Vgl. Küppers 1986, 211). Hierbei setzten sich Merkmalsträger gegenüber ihren Konkurrenten durch, obwohl sie keinen

¹⁸³ Vgl. hierzu auch Walter (1997, 191), der den Zusammenhang zwischen dem Tautologievorwurf gegen das Selektionsprinzip und Keils Kritik aber als weniger direkt darstellt.

Selektionsvorteil haben. Dieses Ergebnis ist auch spieltheoretisch nachvollziehbar, wobei sich mit Sicherheit vorhersagen lässt, dass eine Selektion stattfinden wird, d.h. es bleibt nur eine Sorte von Merkmalsträgern am Ende übrig, aber was sich nicht vorhersagen lässt, ist welche Sorte von Merkmalsträgern sich durchsetzen wird (Vgl. Eigen 1975, 73). Die Gleichheit der Selektionswerte von konkurrierenden Merkmalsträgern kann aber in der freien Natur als „Idealfall“ (Eigen 1975, 74) bzw. als Grenzfall betrachtet werden. Bei ungleichen Selektionswerten (differentieller Reproduktion) findet hingegen eine sogenannte darwinistische Selektion statt. Dass sich hinter Darwins Selektionsprinzip keine Tautologie verbirgt lässt sich sowohl theoretisch als auch experimentell zeigen (Vgl. Eigen 1975, 75). Da Lebewesen sehr komplexe Systeme sind, deren Reproduktionserfolg von dynamischen Faktoren abhängt, besteht ein rein praktisches Problem die Selektionswerte im Vorherein zu berechnen (Vgl. Eigen 1975, 74). D.h. aber auch, dass es zumindest prinzipiell möglich ist festzustellen, ob konkurrierende Lebewesen, die Träger unterschiedlicher vererbbarer Merkmale sind, unterschiedliche Selektionswerte haben und welche Merkmale mit welchem Effekt zu diesen unterschiedlichen Selektionswerten beitragen.

Führt nun also die Beschränkung auf vergangene Selektionsprozesse zum Tautologieeinwand? Nein, keineswegs. Ob „am besten Angepasster“ mehr bedeutet als bloss „Überlebender“ hängt, wie wir gesehen haben, davon ab, ob die konkurrierenden Merkmalsträger unterschiedliche Selektionswerte haben. Dies lässt sich nun aber zumindest prinzipiell auch unter Beschränkung auf vergangene Selektionsprozesse feststellen. Daraus folgt contra Keil: Auch wenn sich „Anpassung“ und „Funktion“ lediglich auf vergangene Selektionsprozesse beziehen, führt dies nicht automatisch zur Tautologie des „survival of the survivor“. Dementsprechend müssen Millikan und Papineau die Beschränkung auf die Vergangenheit nicht aufheben und sind somit auch nicht anfällig für die Kritik von Nagel und Achinstein. Kurzum: Auch das zweite Argument von Keil vermag nicht zu überzeugen.

Fazit:

Keils Einwände können nicht überzeugen. Seine Strategie, Millikan bei zwei Grundbegriffen ihrer Theorie intentionale Anleihen nachzuweisen, scheitert. Der Begriff der Gleichheit, der bei Millikans Definition der reproduktiv etablierten Familie eine Rolle spielt, präsupponiert kein intentionales Vokabular¹⁸⁴. Dasselbe gilt für den Begriff der Eigenfunktion. Bezüglich des Begriffs der Eigenfunktion irrt Keil, wenn er meint, dass dieser sich nicht nur vergangene Prozesse beziehen könne und somit ein irreduzibles (nicht naturalistisches) antizipatorisches Moment enthielte.

¹⁸⁴ Dem stimmt Keil mittlerweile auch selbst zu (persönliche Emailkorrespondenz).

3.2.4 Davies: Troubles for Direct Proper Functions

In diesem Abschnitt geht es darum, einen Einwand von Paul Sheldon Davies gegen Millikans Theorie des Gehalts zu rekonstruieren und anschließend zu diskutieren. Der wesentliche Einwand von Davies ist, dass Millikans Definition von Reproduktion nicht mit dem Prozess der Meiose¹⁸⁵ kompatibel ist (Vgl. Davies 1994). Da in Millikans Theorie der Begriff der Eigenfunktion auf dem Begriff der reproduktiv etablierten Familie erster Ordnung und dieser auf dem Begriff der Reproduktion aufbaut (Vgl. Abschnitt 2.1.1.2, insbesondere Abbildung 6), folgt aus der Inkompatibilität, dass Organe von Lebewesen, die sich sexuell fortpflanzen¹⁸⁶, wie menschliche Herzen oder der Detektor-Schnapp-Mechanismus des Frosches, keine Eigenfunktionen haben. Davies Einwand ist für Millikans Lösung des Problems der Missrepräsentation relevant, weil Millikans Lösung auf dem Begriff der Eigenfunktion gründet. Trifft Davies Einwand zu, dann würde Millikans Theorie nur auf Lebewesen mit asexueller Vermehrung Anwendung finden, die für die Kognitionswissenschaften weniger von Interesse sind.

Rekonstruktion von Davies' Einwand

Davies liefert sein Argument bereits in einer Form, die die argumentative Struktur deutlich macht. Dies erspart uns die semi-formale Rekonstruktion. Wir können Davies einfach zitieren:

„The argument against Millikan's theory or definition has the following form:

- (1) Something has a direct proper function only as a member of a reproductively established family; in particular, a higher-order trait (e.g., the possession of a heart) possesses a direct proper function only as a member of a higher-order family,
- (2) A higher-order trait belongs to a higher-order family only if the trait is produced by items (e.g., genes) belonging to a first-order family, where the first-order family has the direct proper function of producing such higher order traits,
- (3) Two items belong to the same first-order family only if they were produced via reproduction from the same character of the same model(s),
- (4) The first-order item that plausibly produces a human heart is a specific set of genes.

Premises (1)-(3) should be clear from section II above and premise (4) should be obvious. The crux of my argument rests upon the fact that

- (5) For the set of genes that plausibly produces any token human heart, it is biologically impossible that this set both is produced via reproduction and belongs to a first-order family with the familial character of "producing hearts",

¹⁸⁵ „Als Meiose (griech. μείωσις meiosis 'Verminderung', 'Verkleinerung'), auch Reifeteilung oder Reduktionsteilung, wird eine besondere Form der Kernteilung diploider Zellen bezeichnet, bei der die Chromosomenzahl halbiert wird, indem jedem der beiden Tochterkerne je ein homologer Partner eines Chromosomenpaars zugeteilt wird. Im Rahmen eines Kernphasenwechsels ist dies der Übergang zur haploiden Phase. Damit einher geht gewöhnlich eine Rekombination.“ (<http://de.wikipedia.org/wiki/Meiose>, 14:44, 29.03.2015).

¹⁸⁶ Die Meiose ist ein Teilprozess der sexuellen Fortpflanzung.

I will argue directly for premise (5) in this section. The arguments that support (5) can be generalized, for they apply to any reasonably complex phenotypic trait. Indeed, the very arguments given for (5) also show that

(6) For the set of genes that plausibly produces tokens of any complex phenotypic traits (e.g., hearts, lungs, brains, kidneys, etc.), it is biologically impossible that that set both is produced via reproduction and belongs to a first-order family with the familial character of producing that trait.

The damaging conclusion, therefore, is that

(7) No higher-order biological traits (e.g., hearts, lungs, etc.) belong to any higher-order reproductively established family and hence no higher-order biological traits possess any direct proper function.“ (Davies 1994, 368-369)¹⁸⁷

Auf die *Prämissen 1-4* möchte ich nicht näher eingehen. Sie entsprechen im Wesentlichen meiner Darstellung von Millikans Theorie (Vgl. Abschnitte 2.1.1.3.1-2.1.1.3.5)¹⁸⁸.

Entscheidend für Davies Argument ist die *Prämisse 5* (Vgl. auch Vogel 2010, 926), die wir uns daher genauer anschauen müssen. Davies selbst stellt die Prämisse als Dilemma dar:

Entweder wir bestimmen die Menge der Gene so, dass sie Herzen produzieren können, dann sind sie aber keine Reproduktionen in Millikans Sinne;

Oder wir bestimmen die Menge der Gene so, dass sie Reproduktionen in Millikans Sinne sind, dann können sie aber keine Herzen produzieren.

„The general argument for (5) can be formulated as a dilemma. On the one hand, if we characterize the set of genes that plausibly produces any token human heart, the set is either too large or too scattered to satisfy Millikan's definition of reproduction. It is a simple biological fact that such a set cannot be copied across generations. On the other hand, if we characterize the relevant set of genes sufficiently narrowly to satisfy the definition of reproduction, the set is thereby rendered too small or too fragmented to satisfy Millikan's definition of familial character. It is a simple biological fact that such a set is too restricted to ever produce any token human heart and is thus incapable of possessing the familial character "producing hearts."“ (Davies 1994, 369)

Die Darstellung der Prämisse als Dilemma ist zwar ansprechend und Davies' Ausführungen folgen auch formal dieser Struktur, dennoch scheint mir der tatsächliche Argumentationsverlauf von dieser Struktur abzuweichen. Meines Erachtens argumentiert Davies für zwei Behauptungen:

1. Millikans Definition der Reproduktion ist nicht mit dem Prozess der Meiose kompatibel.¹⁸⁹
2. Selbst wenn es gelänge diese Inkompatibilität durch eine Modifikation von Millikans Definition der Reproduktion aufzuheben, dann wären die betreffenden Gene nicht in der Lage ein Herz zu produzieren.

Für die erste Behauptung argumentiert Davies wie folgt:

¹⁸⁷ Für eine Übersetzung siehe Vogel (2010, 925).

¹⁸⁸ Allerdings legt die Prämisse 4 nah, dass Davies glaubt, HOREF-Mitglieder müssten von FOREF-Mitgliedern produziert werden. Dies ist m.E. nicht der Fall, die Mitglieder einer HOREF können auch die Produkte von Mitgliedern einer anderen HOREF sein (Vgl. Abschnitt 2.1.1.3.5, S. 74f), solange am Ende dieser "Produktionskette" Mitgliedern einer FOREF stehen.

¹⁸⁹ Insbesondere dieser erste Punkt entspricht m.E. nicht dem ersten Horn des Dilemmas. Die Rekombination, die aufgrund der Meiose entsteht, macht so oder so Schwierigkeiten, ganz unabhängig davon, ob es sich um zu viele oder zu verteilte Gene (too large or too scattered) handelt. Wir werden auch gleich sehen, dass dies keine Rolle in Davies erster Teilargumentation spielt.

Er betrachtet nacheinander die drei plausibelsten *models*¹⁹⁰ für eine reproduktive etablierte Familie erster Ordnung, deren Mitglieder für die Produktion von Herzen zuständig sein könnten, und zeigt dann jeweils, dass es mit dem *character*¹⁹¹ der reproduktiv etablierten Familien, der sich aus den jeweiligen *models* ergeben müsste, Schwierigkeiten gibt, weil dieser nicht mit dem übereinstimmt, was Millikans Definition von Reproduktion als *character* vorsieht (Vgl. Davies 1994, 370).

Mit anderen Worten: Nach Millikan haben reproduktive etablierte Familien erster Ordnung immer ein Modell. Betrachtet man jedoch die plausibelsten Kandidaten, die als Modelle für die Gruppe von Genen, die Herzen produzieren, in Frage kommen, dann zeigt sich, dass diese Kandidaten nicht Millikans definitorische Ansprüche erfüllen. Dies wiederum läuft darauf hinaus, dass die Gruppe von Genen, die Herzen produzieren, gar keine reproduktive etablierte Familie erster Ordnung ist, denn es gilt: Ohne Modell, keine reproduktive etablierte Familie erster Ordnung¹⁹². Daraus ergibt sich wiederum: Ohne reproduktive etablierte Familie erster Ordnung, keine reproduktive etablierte Familie höherer Ordnung. Ohne reproduktive etablierte Familie höherer Ordnung, keine Eigenfunktion von Organen.

Bevor wir aber Davies Argumentation für die erste Behauptung genau nachvollziehen können, müssen wir noch drei Punkte berücksichtigen.

Die ersten beiden Punkte betreffen Millikans Theorie.

- 1) Millikan bezeichnet dasjenige, wovon etwas eine Reproduktion ist als Modell. Wenn B also die Reproduktion von A ist, dann ist A das Modell von B. Modelle gibt es also folglich genauso wie Reproduktionen immer nur hinsichtlich bestimmter Eigenschaften (Vgl. Abschnitt 2.1.1.3.1).
- 2) Da Modelle selbst nicht mit zur reproduktiv etablierten Familie gehören und Eigenfunktionen erst nach einer Generation *innerhalb* einer reproduktiv etablierten Familie entstehen können, müssen wir hier mindestens drei Generationen betrachten.

Der dritte Punkt betrifft die sexuelle Fortpflanzung.

- 3) Die Gene, die bei mir für die Produktion meines Herzens zuständig waren, stammen teilweise von meiner Mutter und teilweise von meinem Vater. D.h. meine herzproduzierenden Gene sind die Reproduktion eines komplexen Objektes, nämlich einer Konfiguration von Genen, die von meinen Eltern stammen.

Wir können uns den dritten Punkt an einem fiktiven Beispiel veranschaulichen, wobei wir der Einfachheit halber annehmen, dass lediglich zwei Gene (mit beiden Allelen) an der Produktion von Herzen beteiligt sind. Wegen Punkt 2 müssen wir 3 Generationen von herzproduzierenden Genen betrachten. In Abbildung 66 ist nur die väterliche Seite vollständig darstellt.

¹⁹⁰ Gemeint sind hier *models* in Millikans Sinn. Das „Original“ von dem eine Reproduktion abstammt bezeichnet Millikan als Modell (*model*). D.h., wenn B die Reproduktion von A ist, dann ist A das Modell von B (Vgl. Abschnitt 2.1.1.3.1, 52).

¹⁹¹ Gemeint ist hier *character* in Millikans Sinn. Die Gesamtheit der reproduktiv etablierten Eigenschaften bezeichnet Millikan auch als reproduktiv etablierten Charakter oder kurz *character* (Vgl. Abschnitt 2.1.1.3.1, 52).

¹⁹² Vorausgesetzt, dass sich neben den plausiblen Kandidaten nicht noch weitere finden (mögen diese auch unplausibel sein), die dann doch Millikans definitorische Ansprüche erfüllen.

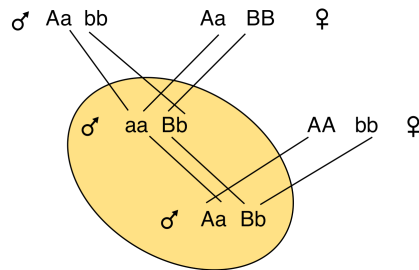


Abbildung 66: Meine herzproduzierenden Gene (Aa Bb) stammen teilweise von meiner Mutter (AA bb), von ihr habe ich A und b bekommen, und teilweise von meinem Vater (aa Bb), von ihm habe ich a und B bekommen. Die kleinstmögliche reproduktiv etablierte Familie von herzproduzierenden Genen müsste die Gene meines Vaters und meine Gene umfassen (gelb unterlegte Fläche)

Um die Überlegungen möglichst einfach zu halten, versucht Davies lediglich seine herzproduzierenden Gene und die seines Vaters in eine reproduktive etablierte Familie erster Ordnung zu bringen (Vgl. Davies 1994, 370, siehe auch Abbildung 66).

Kommen wir also zu Davies Argumentation für die Behauptung, dass die drei plausibelsten *models* für die reproduktive etablierte Familie erster Ordnung, die in Abbildung 66 exemplarisch durch die gelb unterlegte Menge dargestellt ist, zu Schwierigkeiten bezüglich des *characters* dieser Familie führen.

Der **erste Vorschlag**, den Davies betrachtet, lautet, dass das model für die reproduktiv etablierte Familie die Gesamtmenge (conjunction) der herzproduzierenden Gene der Großeltern väterlicherseits ist (Vgl. Davies 1994, 370). Im Beispiel also: Aa BB + Aa bb. Ohne weitere Begründung identifiziert Davies dann den *character* der reproduktiv etablierten Familie mit dessen *model*.

„[...] it seems reasonable to simply identify the character of the family with the model itself.“ (Davies 1994, 370)

D.h. der *character* der reproduktiv etablierten Familie wäre die Eigenschaft, die Menge Aa BB + Aa bb zu sein. Dieser Vorschlag funktioniert aber offensichtlich nicht, denn während der Meiose findet eine Reduktion um die Hälfte der Gesamtmenge statt (Vgl. Davies 1994, 370-371). Folglich können die Mitglieder der reproduktiv etablierten Familie nicht diesen *character* haben. Damit ist nach Davies der erste Vorschlag gescheitert.

Der **zweite Vorschlag**, den Davies betrachtet, geht bereits im vornherein von typen-identischen Gensets bei den Großeltern väterlicherseits aus. In unserem Beispiel hätten also beide Großeltern entweder das Set Aa BB oder das Set Aa bb. Nehmen wir mal an, es wäre Aa BB und nehmen wir weiterhin an, dass während der Meiose die Gensätze *komplementär* weitergegeben werden. D.h., wenn vom Großvater a B kommt, dann kommt von der Großmutter A B. Wäre dies tatsächlich der Fall, dann hätten wir jetzt eine reproduktiv etablierte Familie mit dem Genset Aa BB als *model* und dem *character* das Genset Aa BB zu sein. Dieser Vorschlag scheitert aber schlicht an der Annahme einer komplementären Zuordnung der Gene während der Meiose. Die Reduktion erfolgt in der Realität rein zufällig. D.h. es ist durchaus möglich, dass vom Großvater und der Großmutter die Allele a B kommen und der Vater somit das Genset aa BB erhält. Daraus folgt aber, dass die reproduktiv etablierte Familie keinen einheitlichen *character* hat bzw. genauer, dass es sich gar nicht um reproduktiv etablierte Familie handelt, weil es eine solche nur relativ zu bestimmten Eigenschaften (*character*) geben kann. Damit ist nach Davies auch der zweite Vorschlag gescheitert.

Beim **dritten Vorschlag** geht Davies davon aus, dass wir in der menschlichen Population eine Disjunktion von Gensets isolieren können, die folgende Bedingungen erfüllen:

- a) sie sind genetisch verschieden
- b) jeder Organismus in der Population besitzt mindestens eines dieser Gensets
- c) die Gensets sind in spezifischer Weise kausal relevant für die Produktion von Herzen
- d) die Disjunktion der Gensets ist irgendwie bestimmend für den *character* der reproduktiv etablierten Familie¹⁹³ (Vgl. Davies 1994, 373).

In unserem Beispiel wäre die Disjunktion der Gensets wahrscheinlich Folgende:

AA BB	oder	Aa BB	oder	aa BB	oder
AA Bb	oder	Aa Bb	oder	aa Bb	oder
AA bb	oder	Aa bb	oder	aa bb	

Davies bezweifelt, dass dieser Vorschlag funktioniert, weil er mit zwei Bedingungen von Millikans Definition der Reproduktion nicht vereinbar ist.

Die erste Bedingung, die nach Davies nicht erfüllt wird, ist, dass die reproduktiv etablierten Eigenschaften bei Reproduktionen Determinanten einer Determinablen sein sollen (Vgl. Davies 1994, 373 und auch Abschnitt 2.1.1.3.1). Was sollen aber, so fragt Davies, die identifizierten Gensets gemeinsam haben und welche anderen Eigenschaften werden durch diese gemeinsamen Eigenschaften ausgeschlossen und zwar in einem Sinne, in dem mehrere Objekte hinsichtlich ihrer Farbe (Determinable) rot (Determinate) sein können und damit ihr Grünsein ausgeschlossen ist? (Vgl. Davies 1994, 373) Davies hält es für nahezu ausgeschlossen, dass es auf diese Frage eine positive Antwort gibt, hauptsächlich weil es sich um genetisch verschiedene Sets handelt.

„But the three subsets of genes listed above [Davies bezieht sich auf sein eigenes Beispiel¹⁹⁴] are, by hypothesis, genetically distinct, so they cannot fall under the same determinable by virtue of their brute physical properties. If not these, however, then which? It seems false or perhaps nonsensical to say that they share a common property or a common contrary property.“ (Davies 1994, 373)

Die zweite Bedingung, die nach Davies nicht erfüllt wird, ist die kontrafaktische Behauptung aus Millikans Definition der Reproduktion „Wäre A (das Model) anders gewesen, dann wäre auch B (die Reproduktion) anders“. Davies Begründung ist Folgende: Die elterlichen Gensets könnten anders gewesen sein, ohne dass sich mein Genset zwingend verändert hätte. Dies liegt wiederum daran, dass während der Meiose jeweils nur ein Teil der elterlichen Gene mein Genset bilden. Davies bringt folgendes Beispiel:

„For example, a gene in my mother may have differed, but because the chromosome containing that different gene was not passed on (during random assortment), my subset would have been unaltered. It is possible, in other words, to produce a proper subset identical to my subset even if one or both members of the mother-father pair differ.“ (Davies 1994, 373-374)

Wir können uns dies auch an unserem Beispiel noch einmal veranschaulichen:

¹⁹³ Bei dieser Bedingung ist mir nicht klar, was Davies damit meint. Im Original lautet die Bedingung: “(d) this disjunction is not merely descriptive but somehow or other definitive of family character” (Davies 1994, 373)

¹⁹⁴ Wir können hier aber einfach an drei beliebige Diskunktionsglieder der oben angeführten Disjunktion denken, z.B. AA BB, AaBB, aa Bb.

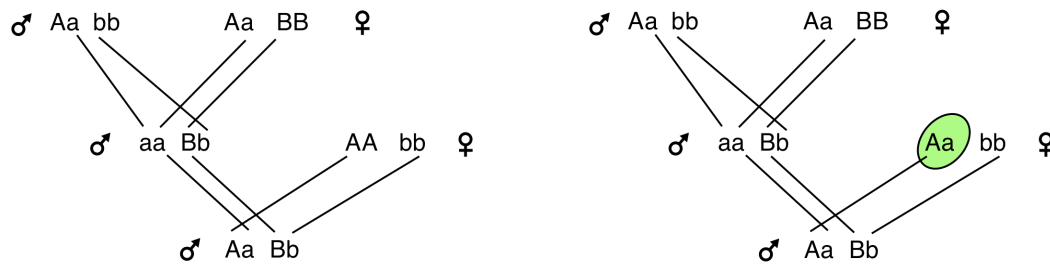


Abbildung 67: Wenn meine Mutter andere Gene gehabt hätte (Aa bb statt AA bb), dann könnte es trotzdem sein, dass dies keinen Einfluss auf mein Genset (Aa Bb) hat.

Der dritte Vorschlag scheitert also nach Davies vor allem, weil er mit Millikans Definition der Reproduktion nicht vereinbar ist. Damit ist für Davies hinreichend gezeigt, dass die Menge der Gene, die Herzen produzieren, keine Reproduktionen in Millikans Sinne sind.

Kommen wir also nun zu Davies Argumentation für die zweite Behauptung, dass selbst wenn diese Inkompatibilität mit dem dritten Vorschlag durch eine Modifikation von Millikans Definition der Reproduktion aufhebbar wäre, dass dann die betreffenden Gensets nicht in der Lage wären ein Herz zu produzieren.

Davies Punkt ist hier einfach der, dass die betreffenden Gensets nur ein Teil dessen sind, was kausal notwendig ist um ein Herz zu produzieren.

„That is, a small, proper subset is but a portion of what is in fact causally necessary; no such subset could produce a human heart in the way that human hearts are typically produced. Hence it is simply false to claim that any such subset has the character of ‚producing hearts.‘“ (Davies 1994, 374)

Dem naheliegenden Einwand, dass jede Aktivität an passende Rahmenbedingungen geknüpft ist, widmet Davies einen eigenen Abschnitt. In diesem Abschnitt versucht Davies zu zeigen, dass die Annahme, wir könnten überhaupt Gensets isolieren, die Herzen produzieren, falsch ist. Die Annahme beruht nämlich auf einer falschen Vorstellung des Zusammenhanges zwischen Genotyp und Phänotyp (Vgl. Davies 1994, 375).

Davies zufolge gibt es zwei Gründe, weswegen wir zu der falschen Vorstellung gelangen, ein kleines Genset wäre kausal herausragend bei der Hervorbringung von phänotypischen Eigenschaften.

Der erste Grund ist unser Wissen, dass bestimmte phänotypische Eigenschaften, wie die Augenfarbe oder die Sichelzellanämie, durch die Veränderung eines einzelnen Gens geändert bzw. eliminiert werden (Vgl. Davies 1994, 377). Aber anders als bei den genannten Eigenschaften ist die Eigenschaft ein Herz zu haben eine Voraussetzung für einen lebensfähigen Organismus. Es ist dementsprechend wenig überraschend, so Davies, dass für die Produktion eines Herzens ein ganzes Ensemble von kooperierenden Genen nötig ist, während die Störung eines Gens ausreicht um eine Änderung wie die Sichelzellanämie hervorzurufen. Mit anderen Worten: Unser Wissen über die phänotypischen Auswirkungen von Mutationen einzelner Gene verfängt nicht im Bezug auf phänotypische Eigenschaften, wie dem Besitz lebenswichtiger Organe.

Der zweite Grund, weswegen wir nach Davies zu der falschen Vorstellung gelangen, ein kleines Genset wäre kausal herausragend bei der Hervorbringung von phänotypischen

Eigenschaften, ist die Entdeckung sogenannter homeotischer Gene. Diese Gene kontrollieren die Entwicklung von ganzen Körperteilen, wie Beinen oder Fühlern¹⁹⁵.

„For example, in the common fruit fly *Drosophila*, if the *Antennapedia* gene is mutated, the fly will develop legs where it normally would develop antennae.“ (Davies 1994, 377)

Die Effekte, die diese Gene haben, wurden in zahlreichen Experimenten nachgewiesen und scheinen sogar stammübergreifend zu sein (Vgl. Carroll 2007, 64-71). Es ist daher versuchend anzunehmen, dass diese Gene eine herausragende kausale Rolle bei der Entwicklung von phänotypischen Eigenschaften spielen (Vgl. Davies 1994, 377). Diese Annahme hält Davies aber für verfehlt. Er begründet dies mit der Tatsache, dass die homeotischen Gene zwar andere Genabschnitte aktivieren können, sie aber ihrerseits auch durch wieder andere Gene aktiviert bzw. gehemmt werden. In diesem regulatorischen Netzwerk von Genen nun eine Gruppe von Genen als kausal herausragend hinzustellen, wäre, so Davies, reine Willkür.

„Homeotic genes and non-homeotic genes alike satisfy important counterfactuals; it is the hierarchically organized system of genes as a whole that is responsible for normal cell differentiation. If one wants to isolate causally special bits within the entire system of genes, one can provide no reason for privileging the homeotic over the non-homeotic. It is therefore simply arbitrary to carve into this system of genes hoping to single out some causally special bit.“ (Davies 1994, 378)

Damit ist für Davies der Gegeneinwand bezüglich der Rahmenbedingungen entkräftet und die Argumentation insgesamt abgeschlossen.

Diskussion von Davies' Argument

Das Argument wurde so rekonstruiert, dass der Schluss aus den Prämissen folgt. Es bleibt also kritisch zu prüfen, ob auch die Prämissen wahr sind. Wie schon bei der Rekonstruktion konzentriere ich mich auch bei der Diskussion auf die fünfte Prämisse.

In meiner Diskussion der fünften Prämisse werde ich zudem die Auseinandersetzung von Matthias Vogel mit Davies Argumentation (Vogel 2010) berücksichtigen. Insgesamt teile ich Vogels Einschätzung, dass Davies Argumentation Ernst zu nehmen, aber nicht vernichtend ist. Im Detail halte ich jedoch einige seiner Entkräftungsversuche für nicht gelungen.

Wie in der Rekonstruktion dargelegt, besteht Davies Rechtfertigung für die fünfte Prämisse im Wesentlichen aus der Argumentation für zwei Behauptungen.

Zur Erinnerung Prämisse 5 lautet:

„For the set of genes that plausibly produces any token human heart, it is biologically impossible that this set both is produced via reproduction and belongs to a first-order family with the familial character of ‚producing hearts‘“ (Davies 1994, 369)

und die beiden Behauptungen für die Davies argumentiert waren:

1. Millikans Definition der Reproduktion ist nicht mit dem Prozess der Meiose kompatibel.

¹⁹⁵ “The discoveries first of Hox genes and of other transcription factors that together make up the regulatory gene networks controlling the expression of genes and the differentiation of embryonic Anlagen and then of the high degrees of conservation of these very genes have received the most attention. [...] Among the ideas that emerged in the context of these investigations are the notion of a genetic toolkit for development – a set of regulatory elements that are involved in the development of the main features of animal bodies, such as segmental patterning and axis formation – and the proposal to reconstruct a so-called Urbilateria as the ancestral condition of all higher animals.” (Laubischler 2007, 350)

2. Selbst wenn es gelänge diese Inkompatibilität durch eine Modifikation von Millikans Definition der Reproduktion aufzuheben, dann wären die betreffenden Gene nicht in der Lage ein Herz zu produzieren.

Betrachten wir also zunächst Davies Argumentation für die erste Behauptung. Diese versuchte Davies zu belegen, indem er nacheinander die drei plausibelsten *models* für eine reproduktive etablierte Familie erster Ordnung, deren Mitglieder für die Produktion von Herzen zuständig sein könnten, betrachtete und dann jeweils zeigte, dass es mit dem *character* der reproduktiv etablierten Familien, der sich aus den jeweiligen *models* ergeben müsste, Schwierigkeiten gibt. Ich konzentriere mich hier auf Davies Überlegungen zum dritten model-Vorschlag, weil ich diese für besonders problematisch halte.

Generell stützt sich Davies Argumentation gegen die drei Vorschläge auf die Annahme, dass der *character* mit dem *model* zu identifizieren ist.

„[...] it seems reasonable to simply identify the character of the family with the model itself.“ (Davies 1994, 370)

Diese Annahme ist ungerechtfertigt und sie entspricht keineswegs Millikans Auffassung. Nach dieser Auffassung müsste eine Reproduktion immer so etwas wie eine Kopie oder ein Duplikat sein. Nach Millikan kann aber z.B. ein Fußabdruck im Sand die Reproduktion eines Fußes (model) sein. Was Davies hier also ungerechtfertigterweise miteinander identifiziert sind die weitergegebenen Eigenschaften (z.B. eine bestimmte Kettenlänge zu haben, ein bestimmtes molekulares Gewicht zu haben) mit dem genetischen Material (einem bestimmten DNA-Abschnitt) selbst. Mit dieser Annahme vereinfacht sich Davies seine Argumentation erheblich.

Gegen den dritten Vorschlag eines *models* für die anvisierte reproduktiv etablierte Familie wendet Davies ein, dass er in zweierlei Hinsicht nicht mit Millikans Definition von Reproduktion vereinbar ist:

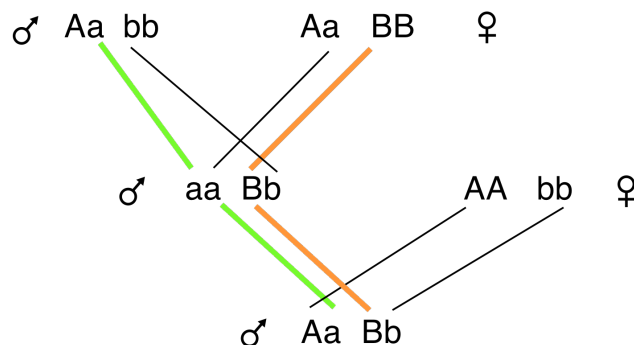
- a) die reproduktiv etablierten Eigenschaften (character) der anvisierten Familie lassen sich nicht als Determinaten einer Determinable verstehen
- b) die kontrafaktische Aussage „Wäre das model anders gewesen, dann wäre auch die Reproduktion anders“ gilt nicht.

Nach Davies ist a) der Fall, weil die reproduktiv etablierten Eigenschaften der Gensets aufgrund ihrer genetischen Verschiedenheit nicht in ihren physikalischen Eigenschaften liegen können. Dies ist m.E. schlichtweg falsch. Zwei verschiedene Basensequenzen, etwa ATGGC und GATCG, können trotz ihrer Verschiedenheit selbstverständlich physikalische Eigenschaften gemeinsam haben, z.B. das molekulare Gewicht *x*. Diese Eigenschaft wäre darüber hinaus auch als Determinate der Determinable „molekulares Gewicht“ zu verstehen. Wenn die Basensequenzen das Gewicht *x* haben, können sie nicht das Gewicht *y* haben. Natürlich bräuchte es eine andere Eigenschaft als die ein bestimmtes molekulares Gewicht haben, die den Gensets gemeinsam sein müsste, wenn diese Eigenschaft für die Produktion von Herzen verantwortlich sein soll. Dass es eine solche Eigenschaft gibt, ist aber nicht prinzipiell ausgeschlossen, wie Davies zu glauben scheint.

Davies hatte b) damit belegt, dass eine Änderung der Gene der Eltern keine Änderung des Gensets des Nachkommen zur Folge hat, solange die geänderten Gene nicht diejenigen sind, die während der Meiose weitergegeben werden. Dies ist zwar korrekt, bedeutet aber nur dann einen Konflikt mit Millikans Definition von Reproduktion, wenn man wie Davies den *character* einfach mit dem *model* identifiziert, was, wie oben bereits erwähnt, keineswegs zwingend ist. Millikans kontrafaktische Bedingung bezieht sich nur auf die reproduktiv etablierten Eigenschaften. Erst durch Davies' Identifikation von *character* und *model*, ergibt sich daraus, dass jede Änderung des *models* auch eine Änderung der Reproduktion nach sich ziehen muss. Dies ist aber nicht prinzipiell der Fall. Eine Änderung des Fußes (model) etwa

hinsichtlich der Hautfarbe muss keine Änderung des Fußabdruckes (Reproduktion) zur Folge haben, eine Änderung der Form hingegen schon. Wenn wir also annehmen, dass es physikalische oder biochemische Eigenschaft gibt, die allen Gensets zukommt und zugleich relevant für die Entwicklung von Herzen ist, dann wird Davies Einwand völlig unproblematisch. Die Veränderung z.B. des mütterlichen Erbteils wäre dann, keine Veränderung hinsichtlich der reproduktiv etablierten Eigenschaft, denn alle Gensets sollen ja diese Eigenschaft haben.

Auch wenn also die beiden Einwände von Davies gegen den dritten Vorschlag nicht greifen, bleibt der Vorschlag dennoch problematisch für Millikan. Betrachten wir noch einmal unser Beispiel.



Wir könnten für jedes Allel die Vorgänger zurückverfolgen (in der Abbildung: oranger Pfad für B und grüner Pfad für a). Die Allele entlang dieser Pfade bilden eine reproductiv etablierte Familie erster Ordnung in Millikans Sinne. Das Problem ist, dass diese reproductiv etablierten Familien nicht diejenigen sind, die Herzen produzieren. Andererseits ist klar, dass ich ein herzproduzierendes Genset habe, weil meine Eltern ein solches haben, und diese haben es, weil meine Großeltern ebenfalls ein solches hatten, usw. Der Zusammenhang zwischen den Gensets der Generationen entspricht also zumindest im Kern dem Zusammenhang in einer reproductiv etablierten Familie erster Ordnung. Millikan müsste also die Definition der Reproduktion so modifizieren, dass sie dem Phänomen der Rekombination Rechnung trägt. Diese Einschätzung teile ich mit Vogel, der im Gegensatz zu mir auch einen Vorschlag hat, wie die Modifikation aussehen könnte (Vgl Vogel 2010, 931).

Vogel schreibt selbst, dass sein Vorschlag noch genauer durchdacht werden müsse und ich muss zugeben, dass ich den Vorschlag nicht so ganz verstehe. Sein Vorschlag lautet jedenfalls:

„Die Idee ist folgende: Kreuzungsfähige Gameten enthalten (infolge von Prozessen wie Crossing-over) ein Gemisch väterlicher und mütterlicher Gene. Wenn wir sagen, dass es für *jedes* dieser Gene (abgesehen von Mutationen) entweder ein väterliches oder ein mütterliches Modell gibt, dann haben wir es zwar mit einer gewaltigen Inflation der Modelle zu tun, dennoch aber nicht mit einer willkürlichen Redeweise. Denn: Wenn wir ein diploid vorliegendes Genom extrem vereinfachend als eine zweiseitige Tabelle betrachten, in der in jeder Zeile zwei dreisrittige Anweisungen notiert sind, dann sorgt der reproductiv etablierte Charakter der Genomfamilie dafür, dass die haploiden Halbtabelle, die die Gameten enthalten, den *gleichen strukturellen Aufbau* haben wie die elterlichen Tabellen. Das, was in Zeile 15 der einspaltigen Tabelle eines Gameten steht, wird durch das bestimmt, was in einer der beiden Spalten entweder des mütterlichen oder väterlichen Genoms in Zeile 15 steht. Zwar setzt sich das Material der einspaltigen Tabelle des Gameten nicht exklusiv aus einer der beiden elterlichen Tabellenspalten zusammen, aber es stammt immer aus einer der beiden Spalten. Es gibt – mit anderen Worten – für jede Zeile des haploiden Gensatzes eines Gameten eine *bestimmbare* Eigenschaft, nämlich die Eigenschaft, eine *mögliche Basensequenz*

(ein Codon) zu enthalten, die entweder durch das väterliche oder mütterliche Material in Zeile n bestimmt wird. Und das gilt, von möglichen Komplikationen abgesehen, für jede weitere Zeile in den Halbtabeln der Gameten.“ (Vogel 2010, 930f)

Neben diesem Vorschlag entwickelt Vogel weitere Ansätze um Davies' Argumentation zu entkräften, die ich für weniger gelungen halte.

Zunächst versteht Vogel Davies' ersten Kritikpunkt folgendermaßen: Die Gene, die in unseren Vorfahren Herzen produziert haben, waren diploid. Die Gene, von denen unsere Gene direkt abstammen, waren jedoch haploid und haben keine Herzen produziert. Die Kette der Reproduktionen von Genen, die Herzen produzieren, ist unterbrochen.

„Das Argument scheint dann zu sein: Es gibt keine transgenerationellen reproduktiven Ketten von diploiden Genmengen, die etwas mit der Produktion von Herzen zu tun haben, weil diese Ketten von haploiden Episoden unterbrochen werden, in denen Genmengen vorliegen, die keine Herzen produzier(t)en. Daher sind diejenigen Gene, die bei unseren Vorfahren Herzen produziert haben, nicht die Modelle unserer (herzproduzierenden) Gene.“ (Vogel 2010, 926)

Dieses m.E. nicht von Davies aufgeworfene Problem versucht Vogel durch das Beispiel eines Kopierers zu entschärfen, der zwei Scanner hat, die nur gemeinsam eine Kopie erstellen können. Vogel sieht aber, dass der eigentliche Einwand von Davies mit der durch die Meiose verursachten Rekombination zusammenhängt und der modifizierte Kopierer hierbei nicht weiterhilft¹⁹⁶.

Weiterhin plädiert Vogel dafür, dass wir nur das gesamte Genom als die eigentliche reproduktiv etablierte Familie erster Ordnung ansehen sollen, deren Eigenfunktion es sein soll Menschen zu „produzieren“ (Vgl. Vogel 2010, 929-930). Dieser Zug scheint mir wenig hilfreich, wenn wir die reproduktiv etablierte Familie erster Ordnung suchen, deren Eigenfunktion es ist Herzen zu produzieren, weil Herzen andernfalls keiner reproduktiv etablierte Familie höherer Ordnung angehören und deshalb keine Eigenfunktion haben können. Im Gegenteil dieser Vorschlag von Vogels verstärkt nur Davies' Einwände.

Insgesamt scheint Vogel zudem die Hoffnung zu haben, man könne Millikans Bemerkung, dass ein Objekt die Reproduktion mehrerer Modelle sein kann, gegen Davies' Einwände fruchtbar machen (Vgl. Vogel 2010, 928). Millikan bringt diesbezüglich das Beispiel eines Schmuckstückes, das zugleich die Nachahmung eines Kreuzes und eines Olivenzweiges ist (Vgl. Millikan 1984, 21). Mir scheint auch dieser Ansatz wenig aussichtsreich, denn ein solches Objekt ist einfach ein Mitglied zweier verschiedener reproduktiv etablierter Familien (Kreuze und Olivenzweige). Was wir aber suchen, ist eine Möglichkeit, bei der die beiden Modelle zusammen mit der Reproduktion *eine* reproduktiv etablierte Familie bilden.

Vogel scheint auch dem zweiten Kritikpunkt von Davies, wonach ein kleines Set von Genen niemals ein Herz produzieren kann, nichts entgegen zu setzen. Im Gegenteil man könnte Vogels Plädoyer dafür, die artspezifischen Genome als einzig wahre reproduktiv etablierte Familie anzusehen, so verstehen, dass er Davies' Kritik zustimmt.

Wir sind auf Davies' zweiten Kritikpunkt bereits unter Abschnitt 2.1.1.3.5 eingegangen. Dort sind wir zu dem Schluss gekommen, dass uns momentan in den allermeisten Fällen keine Normalen Erklärungen zur Verfügung stehen, wie ein bestimmter Teil des Genoms (oder besser wie unter dessen Beteiligung) die ontogenetische Entwicklung einer bestimmten phänotypischen Ausprägung bewirkt.

Davies scheint, wie bereits erwähnt, den Zusammenhang zwischen FOREF und HOREF misszuverstehen, denn er meint, jedes HOREF-Mitglied müsste von einem FOREF-Mitglied

¹⁹⁶ Das Problem sind nicht zwei „Scanner“, sondern zwei Originale, aus denen nur eine „Kopie“ entwickelt wird.

produziert werden. Dies ist nicht der Fall. Nach Millikan kann es HOREFs beliebig hoher Ordnung geben (Vgl. Abschnitt 2.1.1.3.5, 74f). Der kausale Zusammenhang zwischen einem Genset und dessen phänotypischer Auswirkung kann also durchaus sehr indirekt sein. Diese Bemerkung allein reicht aber sicherlich nicht um Davies Kritik zu entkräften.

Insgesamt ist es sehr schwierig Davies' zweiten Kritikpunkt an Millikans Theorie angemessen einzuschätzen, insbesondere, weil sich das Gebiet der Evolutionary Developmental Biology (Kurz: Evo-Devo) in einer rasanten Entwicklung befindet (Vgl. Laubichler 2007, 342). Der Zusammenhang zwischen Genotyp und Phänotyp ist dabei eine der zentralen Forschungsfragen der Evolutionary Developmental Biology (Vgl. Laubichler 2007, 352). Vorerst kann man aber festhalten, dass zumindest die Rede von Genen als Ursachen im Sinne eines spezifischen Unterschiedmachers (specific difference makers)¹⁹⁷ von phänotypischen Merkmalen nach wie vor berechtigt ist (Vgl. Goodfrey Smith 2014, 86-89). Ebenso offensichtlich ist aber auch, dass ein ganzes Netzwerk von Faktoren kausal an der Bildung von komplexen Strukturen, wie Herzen, beteiligt ist. Daraus muss aber nicht streng holistisch folgen, wie Davies zu meinen scheint, dass immer das gesamte Genom an allen ontogenetischen Prozessen beteiligt ist (Vgl. Davies 1994, 372). Eine gewisse Modularität¹⁹⁸ spielt auch bei den ontogenetischen Prozessen eine wichtige Rolle, dies legen zumindest neuere Erkenntnisse der Evolutionary Developmental Biology nahe (Vgl. Kirschner und Gerhard 2007, Kap. 4). Die Mitglieder einer FOREF wären dann keine zusammenhängenden DNA-Abschnitte, sondern ein Netzwerk von verschiedenen lokal verstreuten DNA-Abschnitten. Eine solche FOREF zu identifizieren mag momentan epistemisch unmöglich sein, dies spricht aber nicht prinzipiell gegen eine solche FOREF. Millikans Definition lässt jedenfalls derartige FOREFs zu.

FAZIT:

Der Einwand von Davies lässt sich nicht restlos entkräften. Insbesondere die Inkompatibilität von Millikans Definition der Reproduktion mit den Prozessen der Meiose begrenzt den Anwendungsbereich von Millikans Theorie erheblich. Die kognitiven Leistungen, die Gegenstand der Kognitionswissenschaften sind, stammen zum aller grössten Teil von Lebewesen, die sich sexuell fortpflanzen, und auf diese wäre Millikans Theorie aufgrund der Inkompatibilität nicht anwendbar. Millikans Definition müsste modifiziert werden, um diesen Einwand zu entgehen. Zur Zeit liegt uns m.E. aber kein überzeugender Modifizierungsvorschlag vor.

¹⁹⁷ Zur Anwendung des Begriffs eines specific difference makers in der Genetik (Vgl. Griffith und Stotz 2013, 78-84).

¹⁹⁸ Modularität meint hier einen "Organisationsplan, der sich um halbunabhängige Moduleinheiten von integriertem Design aufbaut. [...] Die Moduleinheiten sind unter Umständen komplex und in sich stark integriert; innerhalb des größeren Plans sind sie jedoch lockerer verknüpft, sodass individuelle Einheiten leicht ersetzt werden können, ohne dass die größere Struktur gefährdet würde." (Kirschner und Gerhard 2007, 385)

3.2.5 Perlman: Conceptual Flux

In diesem Abschnitt geht es darum, eine Kritik von Mark Perlman an Millikans Theorie des Gehalts zu rekonstruieren und anschließend zu diskutieren (Perlman 2000, 137-138). Nach Perlman übersieht Millikan, dass ihre Position mit uneingeschränktem Irrtum kompatibel ist. Millikans Ansatz ist deshalb nicht geeignet zwischen korrekter und inkorrektter Verwendung einer Repräsentation zu unterscheiden, was für eine Lösung des Problems der Missrepräsentation aber nach Perlman eine notwendige Bedingung ist (Vgl. Perlman 2000, 137-138). Da Perlman explizit versucht zu zeigen, dass Millikans Theorie keinen Raum für Missrepräsentationen lässt (Vgl. Perlman 2000, XIX), ist Perlmans Argumentation unmittelbar relevant für diese Arbeit. Perlman hält bereits Fodors Argument (Vgl. Abschnitt 3.2.1) für überzeugend, glaubt aber mit seinem Einwand ein noch stärkeres Argument gegen Millikan vorbringen zu können (Vgl. Perlman 2000, 137).

Rekonstruktion von Perlmans Argument

Perlmans Argument lässt sich meines Erachtens so rekonstruieren:

1. Eine Lösung des Problems der Missrepräsentation (L) muss auf einer Unterscheidung zwischen korrekter und inkorrektter Verwendung (U) einer Repräsentation basieren [Wenn L, dann U]
 2. Nach Millikans Theorie (M) wäre es möglich, dass alle Repräsentationen Missrepräsentationen (A) sind. [Wenn M, dann A]
 3. Wenn alle Repräsentationen möglicherweise Missrepräsentationen (A) sind, dann gibt es keine Unterscheidung zwischen korrekter und inkorrektter Verwendung (U). [Wenn A, dann $\neg U$]
-
4. Millikans Theorie kann das Problem der Missrepräsentation nicht lösen. [Wenn M, dann $\neg L$]

Nachdem wir die Struktur des Argumentes rekonstruiert haben, wollen wir nun die einzelnen Prämissen genauer anschauen.

Die *erste Prämisse* basiert auf einer Überlegung Perlmans, die das Problem der Missrepräsentation bei einfachen Kausaltheorien der Bedeutung (Vgl. Abschnitt 1.3.1) auf Gebrauchstheorien der Bedeutung überträgt: Wenn jeder Gebrauch einer Repräsentation konstitutiv für deren Bedeutung ist, dann ist Missrepräsentation unmöglich. Die Lösung des Problems besteht dann darin zwischen korrektem oder gutem Gebrauch und inkorrektem Gebrauch zu unterscheiden.

„The argument that use theories [...] cannot allow for misrepresentation can be stated informally in the following way: If every actual use of a mental representation or concept is constitutive of its meaning, then there can be no misuse. [...] The most common solution to this problem is to limit the uses of a concept that count in determining its meaning. Rather than say that meaning is a function of all use (leaving no room for misuse), one can say, somewhat imprecisely, that meaning is a function of ‚special’ use, or ‚good’ use.“ (Perlman 2000, 2)

Mit anderen Worten: Ohne eine Unterscheidung zwischen korrektem oder gutem Gebrauch und inkorrektem Gebrauch, gibt es keine Lösung des Problems der Missrepräsentation¹⁹⁹.

Die *zweite Prämisse* besagt, dass es nach Millikans Theorie möglich wäre, dass alle Repräsentationen Missrepräsentationen sind. Diese Annahme ist erläuterungsbedürftig. Zunächst stellt Perlman korrekt fest, dass in Millikans Theorie der Gehalt einer Repräsentation durch die Fälle bestimmt wird, in denen die Repräsentation tatsächlich einen Effekt hatten, der zur Selektion des repräsentationalen Systems beigetragen hat. Diese Fälle interpretiert Perlman dann als Fälle, in denen es der Repräsentation gelingt den relevanten Umweltaspekt vollkommen exakt abzubilden.

„Her [Millikan’s] adaptional account is based on the picture of success being measured by the number of times the function is precisely performed. [...] So, her view of adaptive uses of representations is that they are correct uses, and a cognitive mechanism is adaptive if it yields enough correct uses. [...] The fully adaptive cases are those in which the system gets it exactly right, and this need occur only a small portion of the time for the mechanism to be adaptive and for the organism to survive.“ (Perlman 2000, 137)

Was Millikan dabei übersieht, so Perlman, ist, dass es nach ihrer Sicht nun möglich ist, dass alle Repräsentationen Missrepräsentationen sind. Dies ist möglich, weil adaptiver Erfolg (adaptive success) nicht Exaktheit garantieren kann (Vgl. Perlman 2000, 137). Im Gegenteil, es besteht immer ein Ausbalancieren, ein Tauschgeschäft (trade-off) zwischen der Akkuratheit bzw. Exaktheit auf der einen und Ressourcenverbrauch und Geschwindigkeit auf der anderen Seite (Perlman 2000, 137). Dementsprechend ist es durchaus möglich, dass ein Mechanismus, der schneller und ressourcensparender arbeitet, aber dafür nicht ganz exakt ist, einen Selektionsvorteil gegenüber einem Mechanismus hat, der zwar vollkommen exakt ist, aber zugleich auch sehr langsam oder sehr ressourcenintensiv²⁰⁰. Allgemeiner ausgedrückt: Es gibt keinen direkten Zusammenhang zwischen Adaptivität und Exaktheit. Ein Anstieg der Adaptivität bedeutet nicht zwingend einen Anstieg der Exaktheit und umgekehrt. Es ist also möglich, dass ein Mechanismus selektiert wird, der nie Repräsentationen hervorbringt, die ganz exakt sind.

„The significance of this is that a mechanism could be adaptive even if it never represented things exactly right [...] The fairly accurate system that always gets things approximately right while never getting them exactly right is more adaptive than the system that gets it exactly right some of the time.“ (Perlman 2000, 138)

Den letzten Punkt versucht Perlman durch das Bild zweier Bogenschützen zu illustrieren, wobei der eine immer in die Nähe, aber nie das bulls-eye trifft, während der andere sehr selten das bulls-eye trifft, aber ansonsten weit daneben schießt. Der erste Schütze erzielt somit mehr Punkte als der zweite und würde einen Wettkampf zwischen beiden sicher gewinnen. Wir später auf dieses Bild noch einmal zurückkommen müssen.

Schliesslich ergibt sich daraus, dass es möglich wäre, dass alle Repräsentationen eines repräsentationalen Systems nach Millikan Missrepräsentationen sind, denn wir hatten oben ja festgestellt, dass für Millikan nur die exakten, vollkommen richtigen Repräsentationen keine Missrepräsentationen sind, und es kann Systeme geben, die adaptiv sind und dennoch nie solche exakten, vollkommen richtigen Repräsentationen erzeugen.

¹⁹⁹ Perlman formuliert die Unterscheidung zwischen korrektem und inkorrektem Gebrauch nicht explizit als notwendige Bedingung für eine Lösung des Problems der Missrepräsentation, sondern nur als hinreichende. Eine Interpretation als notwendige Bedingung scheint mir aber nötig, damit sein Argument zwingend ist.

²⁰⁰ Vgl. hierzu auch die Arbeiten von Kahneman und Tversky (Kahneman, Slovic and Tversky 1982).

„So far as adaptive value concerned, Millikan has just got it wrong: it could be adaptive for no representation to get things exactly right, and hence it could be adaptive for them all to be, in Millikan's sense, misrepresentations.“ (Perlman 2000, 138)

Die *vierte Prämisse* besagt, wenn alle Repräsentationen nach Millikans Auffassung möglicherweise Missrepräsentationen sind, dann kann Millikans Auffassung nicht die Unterscheidung zwischen korrekter und inkorrektter Verwendung fundieren. Die Grundidee bei dieser Annahme scheint Folgende zu sein: Wenn alle Repräsentationen nach Millikans Auffassung Missrepräsentationen sind, dann gibt es keine korrekten Repräsentationen und folglich ist eine Unterscheidung zwischen korrekter und inkorrektter Verwendung einer Repräsentation auf der Basis von Millikans Auffassung nicht möglich.

„Thus, adaption [Millikan's view] will not provide the distinction between correct and incorrect representation, since it tolerates total misrepresentation.“ (Perlman 2000, 138)

Diskussion von Perlman's Argument

Das Argument wurde so rekonstruiert, dass der Schluss aus den Prämissen folgt. Es bleibt also kritisch zu prüfen, ob auch die Prämissen wahr sind.

Die *erste Prämisse* scheint mir unbedenklich. Wir wollen sie daher als wahr betrachten und nicht weiter diskutieren.

Die *zweite Prämisse* ist hingegen problematisch. Sie besagt, dass es nach Millikans Theorie möglich wäre, dass alle Repräsentationen Missrepräsentationen sind. Dies ist m.E. schlichtweg falsch.

Zunächst bemerkt Perlman richtig, dass bei Millikan die Fälle des tatsächlichen Erfolgs²⁰¹ für die Bestimmung des Gehalts entscheidend sind. In diesen Fällen macht das repräsentationale System²⁰² per definitionem exakt das, was es machen soll, was seine Funktion ist, und in diesem Sinne macht es in diesen Fällen exakt das Richtige.

Perlman's nächster Überlegungsschritt betraf das Tauschgeschäft (trade-off) zwischen der Akkuratheit bzw. Exaktheit auf der einen und dem Ressourcenverbrauch und der Geschwindigkeit auf der anderen Seite. Der Zusammenhang, den Perlman beschreibt, ist korrekt, dennoch ergibt sich nicht die gewünschte Konsequenz in Bezug auf Millikans Theorie, nämlich, dass alle Repräsentationen bei ihr Missrepräsentationen sein könnten. Machen wir uns dies zunächst an einem Beispiel klar:

Betrachten wir zwei Frösche A und B, in deren Umwelt herumschwirrende kleine, dunkle Objekte zu 50% Fliegen und zu 50% flugfähige Samen²⁰³ sind. Frosch A ist in der Lage zwischen den Samen und Fliegen zu unterscheiden; sein Detektormechanismus verbraucht aber die Hälfte der eingenommenen Ressourcen. Frosch B kann nicht zwischen den Samen und Fliegen unterscheiden; sein Detektormechanismus verbraucht jedoch lediglich ein Fünftel der eingenommenen Ressourcen. Beide Frösche verbrauchen gleich viele Ressourcen für das Schnappen nach den herumschwirrenden Objekten (10 mal Schnappen = 1 Fliege). Wie ein Vergleich über 100 Fälle von vorbeischnappenden Objekten zeigt, ist Frosch B effektiver und würde dementsprechend selektiert werden (Siehe Abbildung 68).

²⁰¹ D.h. Erzeugung eines Effekts, der zur Selektion beigetragen hat.

²⁰² Wenn Perlman von *mechanism* und *system* redet, scheint er aber vor allem den *producer*, also das sensorische System des Frosches im Sinn zu haben.

²⁰³ Die allerdings ungenießbar für die Frösche sind.

	Frosch A	Frosch B
Einnahmen in Fliegen		
Gefangen	50	50
Ausgaben in Fliegen		
Mechanismus	25	10
Schnappen	5	10
Gesamtbilanz	20	30

Abbildung 68: Energiebilanz der Detektor-Schnappmechanismen der Frösche A und B (fiktives Beispiel). Der Mechanismus von B ist effektiver, weil das undifferenzierte Schnappen (-5 gegenüber A) weniger ressourcenintensiv ist als ein differenzierender Detektor (+15 gegenüber A).

Der Detektormechanismus von Frosch B ist weniger exakt aber dennoch effektiver als der Detektormechanismus von Frosch A. Aufgrund der Tauschgeschäftsbeziehung zwischen Exaktheit und Ressourcenverbrauch, wäre es möglich, dass ein Mechanismus selektiert wird, der noch weniger exakt ist als der Detektormechanismus von Frosch B, nämlich wenn er zugleich auch noch viel ressourcensparender ist und dieser Mechanismus könnte einem weniger exakten unterlegen sein, usw.

Was aber nicht sein kann, ist, dass ein Detektormechanismus selektiert wird, der nie dazu führt, dass eine Fliege gefangen wird. Ein Mechanismus, der nie einen positiven Effekt hatte, kann nicht selektiert werden. Genau dies müsste aber der Fall sein, damit nach Millikans Art und Weise der Gehaltsbestimmung alle Repräsentationen Missrepräsentationen sind. Mit anderen Worten: Durch die Art und Weise wie Millikan den Gehalt bestimmt, ist es von vornherein ausgeschlossen, dass alle Repräsentationen Missrepräsentationen sind. Für den Gehalt konstitutiv sind nach Millikan die Fälle, in denen der Mechanismus einen positiven Effekt erzeugt hat. In unserem Beispiel sind dies auch beim Frosch B die Fälle, in denen er die Fliege schnappt. Hätte es für den Frosch B nie Fälle des tatsächlichen Fliegenschnappens gegeben, dann hätte es auch keinen positiven Effekt gegeben und dann wäre der Mechanismus auch nicht selektiert wurden. Wenn es sich also um einen selektierten Mechanismus handelt – und dies ist in Millikans Definition von *intentional icon* sichergestellt²⁰⁴ – dann ist es unmöglich, dass alle Repräsentationen, die der Mechanismus hervorbringt, Missrepräsentationen sind.

Allerdings scheint unser Beispiel einen gewissen quantitativen Aspekt nicht genügend zu berücksichtigen, um den es Perlman aber zu gehen scheint²⁰⁵. Das Bild des Schützen, der mit drei schnellen Schüssen in die Nähe des bulls-eye mehr Trefferpunkt erzielt als ein Schütze, der lange zielt und dann einmal das bulls-eye trifft, verleiht diesem Aspekt Ausdruck.²⁰⁶ Wesentlich ist dabei, dass auch der durchweg nicht exakte Akteur positive Effekte erzeugt, auch der Schnellschütze erzielt ja Trefferpunkte, wenn auch nicht die Maximalanzahl pro Schuss.

Es ist nicht ganz leicht diesen Aspekt auf das Froschbeispiel zu übertragen. Ich möchte es dennoch mit einen wahrscheinlich etwas künstlich anmutenden Beispiel versuchen.

Für den Frosch ist ja nicht nur die Anwesenheit einer Fliege relevant, sondern auch ihre Lokalisation. Nehmen wir an, die Lokalisation ist durch eine Koordinate bestimmt. Eine

²⁰⁴ Vgl. Abschnitt 2.1.1.2.

²⁰⁵ Man könnte auch sagen, dass unser Beispiel nur digitale aber keine analogen Repräsentationen berücksichtigt.

²⁰⁶ Ich habe das Bild von Perlman zu seinen Gunsten modifiziert. Jetzt ist es ein trade-off zwischen Exaktheit und Geschwindigkeit.

völlig exakte Repräsentation enthält die Information, an welcher Koordinate sich eine Fliege befindet und bewirkt ein Ausschnellen der Zunge an die Koordinate. Nehmen wir zusätzlich an, dass das Ausschnellen der Zunge an exakt die Koordinate, an der sich eine Fliege befindet, dazu führt, dass die Fliege mit einer 75%igen Wahrscheinlichkeit an der Zunge kleben bleibt. Schnellst die Zunge lediglich an eine der benachbarten Koordinaten, dann bleibt die Fliege nur mit einer Wahrscheinlichkeit von 25% an der Zunge kleben.

Auch hier wäre es denkbar, dass ein Mechanismus, der *nie* exakt die Koordinaten einer Fliege übermittelt, aber sehr oft nah dran ist, einem Mechanismus überlegen ist, der zwar exakt die Koordinaten übermittelt aber zugleich enorme Ressourcen verbraucht. Zumindest *prima facie* entsteht hier der Eindruck, dass wir es mit einem Mechanismus zu tun haben, der immer missrepräsentiert, weil er nie die exakte Lokalisation der Fliege übermittelt, weswegen die Zunge die Fliege nie optimal erwischt.

Der Eindruck besteht aber nur, wenn wir den Gehalt der Repräsentation – insbesondere den Teil, der die Lokalisation betrifft – durch die Koordinate bestimmen, an die die Zunge schnellst. Der Gehalt wäre also: FLIEGE AN KOORDINATE Y. Dies mag zwar intuitiv plausibel sein, es ist aber nicht die Weise, in der wir nach Millikan den Gehalt der Repräsentation zu bestimmen haben. Das Ausschnellen der Zunge an eine bestimmte Koordinate, die durch die direkte Eigenfunktion der Repräsentation festgelegt ist, hat in der Vergangenheit gelegentlich dazu geführt, dass Fliegen gefangen wurden. Da *ex hypothesi* **nie** die exakte Position der Fliege übermittelt wird, hat sich die Fliege, in den Fällen, in den sie gefangen wurde, an einer Koordinate in unmittelbarer Nachbarschaft zur übermittelten Koordinate befunden (Abbildung 69).

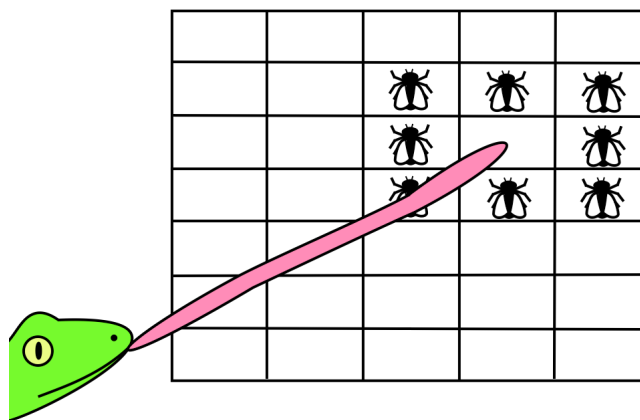


Abbildung 69: Mögliche Positionen, an denen sich eine Fliege bei erfolgreichem Ausschnellen der Zunge befunden hat

Die Normale Bedingung für die direkte Eigenfunktion der Repräsentation x , die dazu führt, dass die Zunge an Koordinate y ausschnellt, ist also, dass sich an einer der benachbarten Koordinaten von y eine Fliege befindet. Der Gehalt der Repräsentation x ist dementsprechend etwas wie: FLIEGE IN UNMITTELBARER NÄHE DER KOORDINATE Y. Daraus wiederum folgt, dass die Repräsentation nicht generell eine Missrepräsentation ist, auch wenn sie niemals die Koordinate übermittelt, an der sich eine Fliege befindet, und der Frosch folglich nie eine Fliege optimal erwischt. Wenn sich tatsächlich in unmittelbarer Nähe der Koordinate y eine Fliege befindet, dann ist die Repräsentation keine Missrepräsentation, sondern korrekt. Entscheidend dafür, dass Perlmans Annahme, Millikans Theorie lasse generelle Missrepräsentation zu, falsch ist, ist die Unvereinbarkeit dieser Annahme mit

Millikans Art und Weise der Gehaltsbestimmung. Jede Repräsentation als Typ hat bei Millikan Normale Bedingungen, was nichts anderes bedeutet als, dass es bereits Repräsentationen (tokens dieses Typs) gegeben hat, die keine Missrepräsentationen waren. Perlmans Irrtum beruht m.E. zu einem grossen Teil auf einem Missverständnis. Exakt richtig machen (get things exactly right) bedeutet bei Millikan, den Effekt erzeugen, der für die Selektion des jeweiligen Funktionsträgers verantwortlich war. Der Effekt kann der maximale oder optimale Effekt dessen sein, was in den betreffenden Fällen erreicht werden kann, aber muss es nicht. Es ist sogar viel wahrscheinlicher, dass es sich nicht um den maximalen oder optimalen Effekt handelt. Perlman versteht Millikan aber so, dass „exakt richtig machen (get things exactly right)“ auf einen maximalen oder optimalen Effekt hinausläuft, was einfach unzutreffend ist.

Damit kann die Argumentation von Perlman gegen Millikan als gescheitert betrachtet werden. Wir wollen aber der Vollständigkeit halber auch noch die dritte Prämisse diskutieren. Auch diese ist fragwürdig. Selbst wenn es Millikans Theorie zulassen würde, dass alle Repräsentationen Missrepräsentationen sind, folgt daraus nicht, dass Millikans Theorie generell ungeeignet wäre zwischen korrektem und inkorrektem Gebrauch zu unterscheiden. Es könnte dann immer noch der Fall sein, dass es über viele Spezies und Wahrnehmungssysteme hinweg, sowohl korrekte als auch inkorrekte Repräsentationen (jetzt in Perlmans Sinn) gibt, obwohl es theoretisch möglich gewesen wäre, dass alle inkorrekt sind. Mit anderen Worten: Für die starke Behauptung, Millikans Theorie könne nicht zwischen korrektem und inkorrektem Gebrauch unterscheiden, müsste Perlman nicht nur zeigen, dass Adaption und Exaktheit nicht unbedingt miteinander verbunden sind, sondern dass sie sich ausschliessen. Eine Argumentation in diese Richtung – etwa: absolute Exaktheit ist generell zu ressourcenintensiv und daher nie adaptiv – ist denkbar, wird aber von Perlman nicht ausgeführt. Strenggenommen würde aber noch nicht einmal dies ausreichen, denn die blossen Möglichkeit, dass alle Repräsentationen falsch sind, ist irrelevant, solange klar ist, unter welchen Bedingungen sie korrekt wären.

FAZIT:

Perlmans Einwand kann nicht überzeugen, weil er auf einem Missverständnis beruht. Perlmans Prämisse, dass bei Millikan alle Repräsentationen Missrepräsentationen sein könnten, basiert auf der falschen Annahme, dass bei Millikan der selektierte Effekt (die Eigenfunktion) immer der optimale oder maximale Effekt ist, der in den betreffenden Situationen erreicht werden kann.

3.2.6 Mendola: Against Etiological Semantics²⁰⁷

In diesem Abschnitt geht es darum, einen Einwand von Joseph Mendola gegen Millikans Theorie des Gehalts zu rekonstruieren und anschließend zu diskutieren. Mendola argumentiert dafür, dass das eigentliche Problem von Millikans Theorie, wie anderen etiologischen Ansätzen, nicht das Disjunktionsproblem (Vgl. Abschnitt 1.3.2) ist, sondern, dass sie den falschen Gehalt liefern, wie sich am Fall der Farbenblindheit zeigen lässt (Vgl. Mendola 2008, 103-130). Mendolas Einwand ist für die vorliegende Arbeit von besonderem Interesse, weil für seine Argumentation deformierte bzw. missgebildete Funktionsträger und deren Auswirkung auf die Gehaltsbestimmung bei Millikans Theorie eine zentrale Rolle spielen. Mendolas Argumentation ist damit die einzige mir bekannte, die den Ansatz, das Problem der Missrepräsentation über deformierte bzw. missgebildete Funktionsträger zu lösen (Vgl. Abschnitt 2.2.3), direkt in Frage stellt.

Rekonstruktion von Mendolas Argument

Der Kern von Mendolas Kritik lässt sich meines Erachtens als *modus tollens* rekonstruieren:

1. Nach Millikans Theorie behalten auch deformierte Funktionsträger ihre Funktion [A1]
2. Nach Millikans Theorie wird der Gehalt durch die Funktion des Gehaltsträgers bestimmt [A2]
3. Nach Millikans Theorie wäre der Gehalt der internen Repräsentation, die der visuellen Wahrnehmungserfahrung entspricht, bei normalen und farbblienden Menschen gleich [Wenn (A1 und A2), dann G]
4. Der Gehalt der internen Repräsentation, die der visuellen Wahrnehmungserfahrung entspricht, ist bei normalen und farbblienden Menschen verschieden. [¬G]
5. Millikans Theorie ist nicht korrekt. [¬(A1 und A2)]

Nachdem wir die Struktur des Argumentes rekonstruiert haben, wollen wir nun die einzelnen Prämissen genauer anschauen.

Die *erste Prämisse* gibt lediglich Millikans Auffassung korrekt wieder, die wir in Abschnitt 2.2.3.1 ausführlich behandelt haben.

Ähnliches gilt für die *zweite Prämisse*. Mendola erfasst hier den Kern von Millikans Gehaltsbestimmung, auch wenn etliche Details unberücksichtigt bleiben (Vgl. die Abschnitte unter 2.1.1.3, die meiner Darstellung von Millikans Gehaltsbestimmung in Abschnitt 2.1.2 vorangegangen sind).

„The content of an inner representational state is, more or less, whatever it is its function or purpose to be caused by, where that function or purpose is crucially constituted by the evolutionary history of the creature with that state [...]“ (Mendola 2008, 115)

Die *vierte Prämisse*, die wir aus Darstellungsgründen vorziehen, besagt, dass sich der Gehalt der internen Repräsentation, die der visuellen Wahrnehmungserfahrung entspricht, bei

²⁰⁷ Der Text basiert auf einen Vortrag, den ich 2013 in Bochum 2013 gehalten habe.

normalen und farblinden Menschen unterschiedlich ist. Worauf Mendola damit hinaus möchte, können wir uns an folgender Überlegung klar machen.

Wenn eine farbenblinde und eine normalsichtige Person mit einem grünen Objekt mit roten Punkten konfrontiert werden, dann unterscheiden sich ihre jeweiligen Wahrnehmungseindrücke und beide Personen würden dementsprechend unterschiedliche Beschreibungen ihrer visuellen Erfahrung abgeben. Stark schematisiert, lässt sich diese Situation, wie folgt, darstellen:

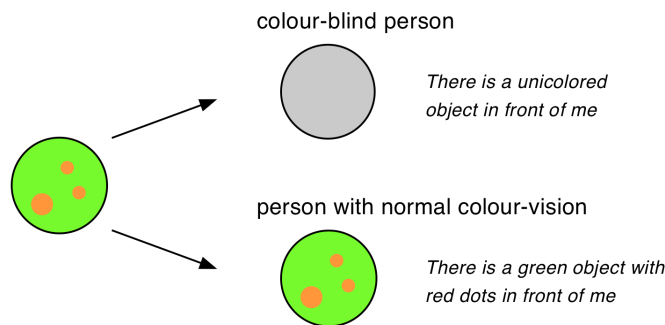


Abbildung 70: Schematisierte Darstellung der Wahrnehmungseindrücke und deren Beschreibung bei einer farblinden und einer normalsichtigen Person, die mit einem grün-rotem Objekt konfrontiert sind

Wenn wir zudem annehmen, dass den Wahrnehmungserfahrungen (sensory states) interne Repräsentationen entsprechen, die einen korrespondierenden Gehalt haben, dann müssen die beiden Personen interne Repräsentationen unterschiedlichen Gehalts haben.

Auch wenn Millikan dies anders sieht, so Mendola, müssen sich etiologische Theorien daran messen lassen, inwiefern, der von ihnen gelieferte Gehalt mit dem phänomenologischen, erfahrenen Gehalt übereinstimmt (Vgl. Mendola 2008, 112).

„What we need here is a match between introspectible human sensory content and the kind that some version of ET can deliver in a properly principled way.“ (Mendola 2008, 112-113)

Kommen wir nun zur *dritten Prämisse*, die für Mendolas Argumentation entscheidend ist. Die Prämisse besagt, dass der Gehalt der internen Repräsentation nach Millikans Theorie bei normalen und farblinden Menschen gleich wäre. Mendola scheint dabei in etwa folgenden Gedankengang zu haben: Die interne Repräsentation der farblinden Person ist zwar deformiert²⁰⁸, aber dies ändert nichts an ihrer Funktion (aus Prämisse 1). Dementsprechend haben die internen Repräsentationen beider Personen dieselbe Funktion. Und weil die Funktion den Gehalt bestimmt (aus Prämisse 2), müssen beide Repräsentationen auch denselben Gehalt haben. Wie bereits bei der Erläuterung der vierten Prämisse vorweggenommen, glaubt Mendola, dass es unplausibel ist, dass die farbenblinde Person eine interne Repräsentation hat, deren Gehalt rot und grün unterscheidet, während die Person davon selbst nichts bemerkt (siehe Abbildung 70).

„Consider again the crucial case of the color blind. If a red-green color blind person can be said to retain systems of production or consumption that yet are deformed from their normal state, systems that evolved to allow certain discriminations by the

²⁰⁸ Genaugenommen ist sie nur das Produkt eines deformierten visuellen Systems und nicht selbst deformiert. Sofern man aber, wie vermutlich Mendola, annimmt, dass die Funktion des visuellen Systems auch die Funktion der Repräsentation bestimmt, läuft beides – deformiert sein oder Produkt eines deformierten Systems zu sein – auf dasselbe hinaus.

differential production or consumption of distinct sorts of representations, representations that yet in fact, we will presume, introduce no differences that the person in question can notice, it is very implausible to claim that nevertheless these differences are present unnoticed in that person's experience.“ (Mendola 2008, 127)

Diskussion von Mendolas Argument

Formal gesehen ist das Argument von Mendola korrekt, d.h. der Schluss folgt aus den Prämissen. Es bleibt also kritisch zu prüfen, ob auch die Prämissen wahr sind.

Die *erste Prämisse* scheint mir unbedenklich und soll deshalb nicht weiter diskutiert werden.

Die *zweite Prämisse* ist zwar nicht falsch, aber sie unterschlägt zu viele Details, die in Millikans Auffassung der Gehaltbestimmung eine wichtige Rolle spielen. Dies führt dann in der *dritten Prämisse* zu dem Resultat, dass die farbenblinde und die normalsichtige Person in der oben geschilderten Situation interne Repräsentationen mit demselben Gehalt haben. Zu diesem Ergebnis gelangt man nicht, wenn man Millikans Theorie genau anwendet (Vgl. Abschnitt 2.1.2). Wir können uns dies auch an einem stark vereinfachten Beispiel, das dem von Abschnitt 2.1.1.1 ganz analog ist, klar machen.

Gehen wir von einer fiktiven Spezies aus, die zur Farbwahrnehmung in der Lage ist, insbesondere, rot von grün unterscheiden kann. Betrachten wir zunächst ein Individuum mit normalem Farbsehvermögen:

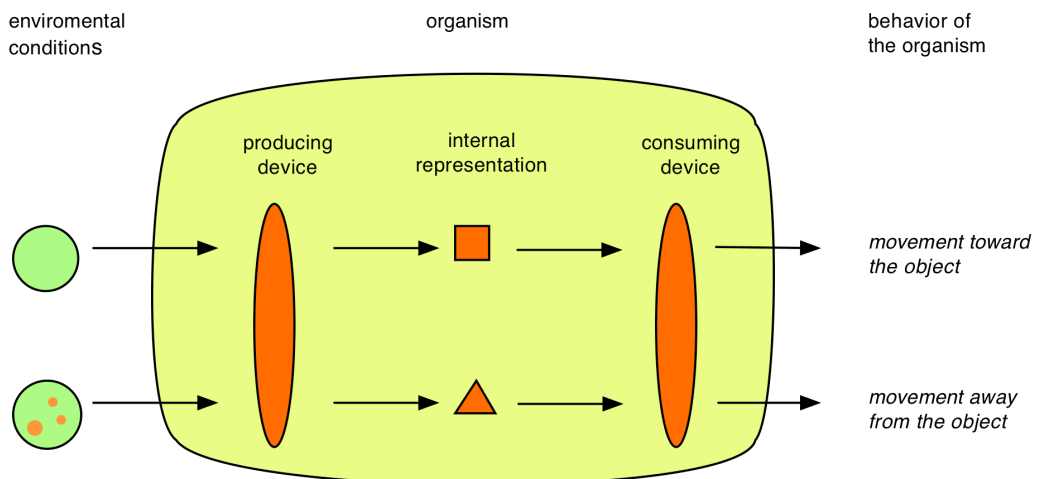


Abbildung 71: Normalsichtiger Organismus

Nehmen wir weiterhin an, dass einfarbige Objekte in der Umwelt des Organismus häufig Nahrung sind, während gepunktete Objekte häufig gefährlich sind.

Nach Millikan hat nun der *producing device* die relationale Eigenfunktion verschiedene *internal representations* (wieder symbolisiert durch ein Viereck und ein Dreieck) in Abhängigkeit von verschiedenen Umweltsituationen (einfarbiges Objekt anwesend und gepunktetes Objekt anwesend) zu produzieren.

In einem konkreten Fall erwirbt der *producing device* eine *adapted proper function*. Sagen wir, ein gepunktetes Objekt befindet sich in Front des Organismus. In diesem Fall erwirbt er die Funktion die durch das Dreieck symbolisierte interne Repräsentation zu produzieren. Dasjenige, woran der *producing device* angepasst (adapted) ist – das gepunktete Objekt –, bezeichnet Millikan als seinen *adaptor* und die interne Repräsentation als einen *adapted device* (Vgl. Abschnitt 2.1.1.6).

Der adapted device, also die durch das Dreieck symbolisierte interne Repräsentation, hat nun zugleich eine abgeleitete Eigenfunktion (derived proper function) und eine direkte Eigenfunktion (direct proper function). Die abgeleitete Eigenfunktion ist vom *producing device* abgeleitet, dessen Funktion über die Produktion der durch das Dreieck symbolisierten internen Repräsentation hinaus, darin besteht, den *consuming device* an die Umweltbedingung, gepunktetes Objekt anwesend, anzupassen und dadurch den *consuming device* dazu zu veranlassen, eine Bewegung vom Objekt weg zu bewirken. Die direkte Eigenfunktion hingegen besteht darin, den *consuming device* an die eigene „Form“ anzupassen und dadurch eine Bewegung vom Objekt weg zu bewirken. Weil die „Form“ der internen Repräsentation aber Normalerweise von einer bestimmten Umweltbedingung abhängt, ist ein *consuming device* der direkt an die „Form“ (symbolisiert durch das Dreieck) angepasst ist auch indirekt an die Umweltbedingung (gepunktetes Objekt anwesend) angepasst. Wenn alles Normal läuft, dann sind die abgeleitete Eigenfunktion und die direkte Eigenfunktion der internen Repräsentationen nicht im Konflikt miteinander. D.h. sie passen den *consuming device* so an, dass eine motorische Reaktion erzeugt wird, die zur gegenwärtigen Umweltsituation passt, also zum Überleben des Organismus beiträgt.

Wenn nun aber der *producing device* wie im Fall der Farbenblindheit deformiert oder missgebildet ist, dann kann es zum Konflikt zwischen der abgeleiteten und der direkten Eigenfunktion der internen Repräsentation kommen. Aufgrund der Deformation des *producing device* soll nun folgendes Schema für den Organismus gelten:

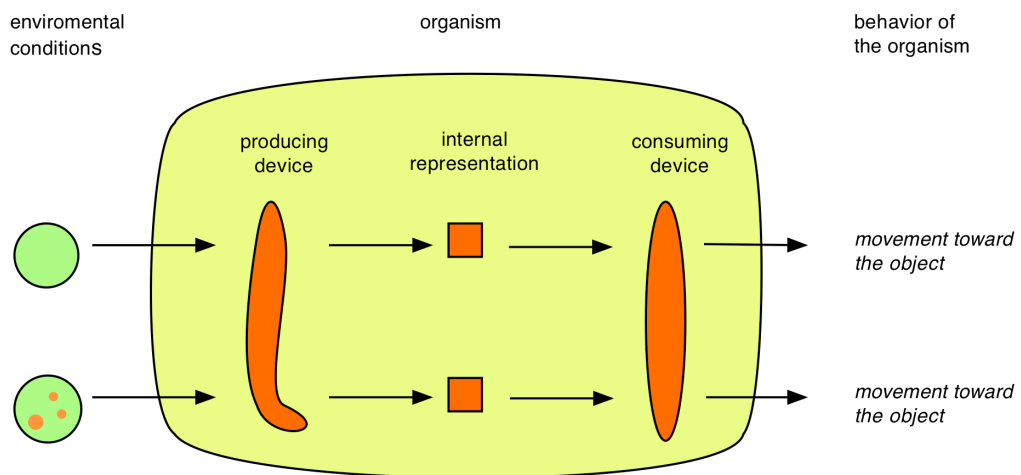


Abbildung 72: Farbenblinder Organismus mit deformiertem producing device

Der *producing device* ist also nicht mehr in der Lage seine Eigenfunktion vollständig zu erfüllen. Konfrontiert mit einem gepunkteten Objekt produziert er eine durch das Viereck symbolisierte interne Repräsentation anstatt einer durch das Dreieck symbolisierten. Nun

haben wir einen Konflikt zwischen der abgeleiteten und der direkten Eigenfunktion. Die abgeleitete Eigenfunktion ist es den *consuming device* dazu zu veranlassen, eine Bewegung vom Objekt weg zu bewirken. Die direkte Eigenfunktion hingegen ist es den *consuming device* an die eigene „Form“ anzupassen, den *consuming device* also dazu zu veranlassen, eine Bewegung zum Objekt hin zu bewirken. Bezüglich dieses Konfliktes schreibt Millikan ausdrücklich:

“[...] the most dominant notion of what is signed by signs is derived by reference to direct proper functions of these signs themselves, hence to resulting adapted proper functions of interpreting devices qua taking these signs as immediate adaptors. It is not derived by reference to adapted functions of the sign’s producing devices.” (Millikan 1984, 43)²⁰⁹

Wenn wir nun „what is signed by signs“ als den Gehalt einer Repräsentation verstehen und die Produkte des *producing device* als Wahrnehmungserfahrungen auffassen, dann unterscheiden sich die Wahrnehmungserfahrungen einer farbenblinden und einer normalsichtigen Person nach Millikans Theorie in Übereinstimmung mit unserem obigen Beispiel. Die farbenblinde Person hat eine Erfahrung mit dem Gehalt „einfarbiges Objekt anwesend“ und die normalsichtige Person hat eine Erfahrung mit dem Gehalt „gepunktetes Objekt anwesend“. Daraus folgt, dass Mendolas Argument an einer Fehlinterpretation von Millikans Theorie scheitert.

Der Vollständigkeit halber wollen wir auch noch kurz auf die vierte Prämisse eingehen. Diese Prämisse ist ebenfalls fragwürdig. Unstrittig ist sicherlich, dass die visuelle Wahrnehmungserfahrung der farbenblinden und der normalsichtigen Person verschieden sind. Problematisch ist hingegen der Zusammenhang zwischen dem Gehalt der internen Repräsentation und dem Gehalt der bewussten Wahrnehmungserfahrung. Mendola verfolgt in seiner Arbeit spezifische Interessen²¹⁰ weswegen seine Gleichsetzung der beiden Gehalte verständlich ist. Wir haben aber bereits bei der Diskussion von Pietroskis Einwand gesehen, dass vor allem im Rahmen dieser Untersuchung von Millikans Theorie die Gleichsetzung der beiden Gehalte eher skeptisch zu sehen ist (Vgl. Abschnitt 3.2.2).

FAZIT:

Mendolas Einwand kann nicht überzeugen. Der Einwand beruht auf einer Fehlinterpretation von Millikans Weise den Gehalt eines Gehaltsträgers zu bestimmen. Insbesondere Millikans Ausführungen zu reationalen Eigenfunktionen werden von Mendola übersehen oder fälschlicherweise als nicht relevant eingeschätzt.

²⁰⁹ Vgl. auch Abschnitt 2.2.3.2.

²¹⁰ “And in any case, my focus in this book is on human thoughts, and among sensory contents merely on the kinds that we can introspect from our armchairs.” (Mendola 2008, 112).

3.2.7 Stegmann: Teleosemantics for Animal Signals

In diesem Abschnitt geht es darum, einen Einwand von Ulrich Stegmann gegen Millikans Theorie des Gehalts zu rekonstruieren und anschließend zu diskutieren. Stegmann stellt fest, dass Millikans Theorie den Gehalt von Tiersignalen (animal signals) nicht generell erklären kann, weil Tiersignale, die der Täuschung dienen, von Millikans Theorie nicht erfasst werden (Vgl. Stegmann 2009). Stegmann hat diesen Einwand bereits in Stegmann (2005) vorgebracht und in Stegmann (2009) weiter ausgearbeitet. Ich konzentriere mich hier auf Stegmann (2009). Stegmanns Einwand steht insofern in unmittelbarem Zusammenhang mit dem Problem der Missrepräsentation als gerade Missinformation durch Täuschung, wie Lügen, als Paradebeispiel einer Missrepräsentation gelten kann. Von einer gelungenen Lösung des Problems der Missrepräsentation könnte man daher erwarten, dass sie zumindest die einschlägigen Formen der Missrepräsentation erfasst²¹¹. Genau dies stellt Stegmann aber in Bezug auf Millikans Theorie in Frage.

Rekonstruktion von Stegmanns Einwand

Der Kern von Stegmanns Kritik lässt sich meines Erachtens als *modus ponens* rekonstruieren:

1. Eine Theorie erklärt den Gehalt von Tiersignalen nicht befriedigend ($\neg B(x)$), wenn diese Theorie einigen Tiersignalen keinen Gehalt zuordnen kann ($\neg A(x)$). [wenn $\neg A(x)$, dann $\neg B(x)$]
2. Die Lichtmuster, die *Photuris versicolor*²¹² aussendet, sind Tiersignale und sie haben einen Gehalt.
3. Den Lichtmustern von *Photuris versicolor* kann nach Millikans Theorie kein Gehalt zugeordnet werden, weil sie keine *intentional icons* sind.
4. Aus 2 und 3: Millikans Theorie ordnet einigen Tiersignalen keinen Gehalt zu [$\neg A(a)$]
5. Aus 1 und 4: Millikans Theorie erklärt den Gehalt von Tiersignalen nicht befriedigend [$\neg B(a)$]

Nachdem wir die Struktur des Argumentes rekonstruiert haben, wollen wir nun die einzelnen Prämissen genauer anschauen.

Die *erste Prämisse* ist sicherlich plausibel, auch wenn eine Theorie nicht gleich nutzlos wird, sobald sie einige Ausnahmen nicht erfasst. Sollte eine Theorie aber häufig auftretende Phänomene in ihrem intendierten Bereich nicht erfassen, dann müssen Alternativen entwickelt bzw. Modifikationen an der Theorie vorgenommen werden. Kurz: Die Theorie kann nicht so bleiben, wie sie ist, weil sie in der gegenwärtigen Form unbefriedigend ist.

Kommen wir zur *zweiten Prämisse*. Um diese später beurteilen zu können, müssen wir das von Stegmann sehr geschickt ausgewählte Beispiel der Lichtmuster von *Photuris versicolor* genauer anschauen. Zunächst beschreibt Stegmann das Partnersuchverhalten von Leuchtkäfern, insbesondere der Art *Photinus macdermotti*.

²¹¹ Täuschungen und Lügen sind aber insofern spezielle Missrepräsentationen als sie vom Produzenten der Missrepräsentation bezweckt werden.

²¹² *Photuris versicolor* ist eine Leuchtkäferart. Siehe auch unten bei der Erläuterung der Prämissen.

„Fireflies are a group of closely related beetle species that use their luminescent abdomina to emit light flashes in order to attract and locate mating partners (e.g., Lloyd 1975). The flashes’ physical properties (e.g., their duration and intensity) are both sex and species specific. Males fly about in their habitat, emitting their typical light flashes. Stationary females respond with their own specific flash if they are willing to mate. A dialogue of flash signals ensues, eventually guiding the male to the female. The signals are taken to convey information about the sex and the species of the sender and its willingness to mate. The flash of a *Photinus macdermotti* male, for instance, is taken to mean something like “here is a *P. macdermotti* male ready to mate.”“ (Stegmann 2009, 866)

Mit diesem Beispiel von Tiersignalen kommt Millikans Theorie gut zurecht. Nach ihrer Theorie besteht die Funktion der Licht emittierenden Mechanismen darin, Lichtmuster zu produzieren, denen bestimmte Eigenschaften (Geschlecht des Senders, Artzugehörigkeit, Bereitschaft zur Paarung) zugeordnet sind. So ist es die Funktion des Licht emittierenden Mechanismus’ eines *P. macdermotti* Weibchens ein bestimmtes Lichtmuster zu produzieren, wenn das Weibchen paarungsbereit ist und das Lichtmuster eines männlichen Artgenossen wahrgenommen wurde. Der die Lichtmuster aufnehmende Mechanismus des Männchens hat dann die Funktion das Männchen zum paarungsbereiten Weibchen zu navigieren. Das Verhalten führt meist zu einer Paarung, was zur Reproduktion der beiden Mechanismen führt (Vgl. Stegmann 2009, 866-867). Stegmann betont, dass es für Millikans Theorie zwingend ist, dass die produzierenden und die konsumierenden Mechanismen ihre Funktion nur kooperativ erfüllen können. D.h. der Erfolg des einen Mechanismus ist vom Erfolg des anderen abhängig und umgekehrt.

„This is Millikan’s technical notion of cooperation: the two devices depend on each other’s success in order to achieve their respective functions in the historically predominant way.“ (Stegmann 2009, 868)

Kommen wir nun zum eigentlich wichtigen Beispiel, dem Verhalten von *Photuris versicolor*. *Photuris versicolor* ist eine Leuchtkäferart deren Beute andere Leuchtkäfer sind, z.B. *P. macdermotti*. Um ihre Beute anzulocken ahmt *Photuris versicolor* die artspezifischen Paarungssignale ihrer Beute nach. D.h. auf das Lichtmuster eines *P. macdermotti* Männchens reagiert sie mit dem spezifischen Lichtmuster, das ein paarungsbereites *P. macdermotti* Weibchen aussenden würde (Vgl. Stegmann 2009, 868). Die Lichtmuster, die *Photuris versicolor* aussendet, haben nach Stegmann zumindest aus verhaltensbiologischer Sicht einen Gehalt und zwar denselben wie die Signale, die sie nachahmen sollen.

„From an ethological point of view, the predator’s female-macdermotti-type flash carries the information that there is a female *P. macdermotti* willing to mate. In other words, the flashes of both a female *P. macdermotti* and its mimic have representational content and the same one, at that.“ (Stegmann 2009, 868)

Dementsprechend sind die Lichtmuster von *Photuris versicolor* ethologisch gesehen Tiersignale.

In der *dritten Prämisse* stecken zwei implizite Annahmen. Zunächst scheint Stegmann stillschweigend davon auszugehen, dass die Lichtmuster von *Photuris versicolor*, sofern sie in Millikans Theorie erfasst werden können, unter Millikans Definition von *intentional icons* fallen müssen (1. Implizite Annahme). Stegmann kann dann zeigen, dass die Lichtmuster von *Photuris versicolor* den Bedingungen eines *intentional icons* nicht genügen, weil sowohl die zweite Bedingung als auch die dritte Bedingung von Millikans Definition nicht erfüllt sind. Hier noch einmal zur Erinnerung die beiden Bedingungen:

„(2) Normally a [intentional icon] stands midway between two cooperating devices, a producer device and an interpreter device, which are designed or standardized to fit

one another, the presence and cooperation of each being a Normal condition for the proper performance of the other.

(3) Normally the [intentional icon] serves to adapt the cooperating interpreter device to conditions such that proper functions of that device can be performed under those conditions.“ (Millikan 1984, 96f)²¹³

Die zweite Bedingung ist nicht erfüllt, weil das Funktionieren des Licht emittierenden Mechanismus' des Jägers definitiv keine Normale Bedingung für den Mechanismus der Beute ist, der das Lichtmuster konsumiert. Ebenso wenig wird der konsumierende Mechanismus der Beute Normalerweise durch das Lichtmuster des Jägers an eine Bedingung angepasst, die es dem Mechanismus der Beute erlaubt seine Eigenfunktion auszuüben, wie dies von der dritten Bedingung gefordert wird.

„Recall one of the elements of cooperation: the normal condition for the consuming device to achieve its function is that the producing device achieves its function. The function of the predator's producing device is to token female-macdermotti-type flashes that map to the presence of a hungry predator (in response to perceiving the flash of a male *P. macdermotti*). The device achieves its function when its flash indeed co-occurs with these circumstances. But the co-occurrence of a female-macdermotti-type flash with the presence of a hungry predator is clearly not the normal condition for the male's consuming device to achieve its function. When males responded in the past to flashes emitted by predators, they typically located a predator instead of a conspecific female. The co-occurrence of a female-macdermotti-type flash with the presence of a hungry predator is therefore not the kind of circumstance that obtained when the consuming device achieved its function (let alone when it achieved its function by the predominant mechanism).“ (Stegmann 2009, 868)

Zudem nimmt Stegmann stillschweigend an, dass nach Millikans Theorie nur dann etwas einen repräsentationalen Gehalt haben kann, wenn es mindestens Millikans Definition von *intentional icon* genügt (2. Implizite Annahme). Deswegen schlussfolgert Stegmann aus der Tatsache, dass die Lichtmuster von *Photuris versicolor* keine *intentional icons* sind, dass die Lichtmuster von *Photuris versicolor* nach Millikans Theorie auch keinen Gehalt haben.

„In other words, the predators' flashes are not of a type where producing devices cooperate with consuming devices in the required sense. They therefore lack representational content on Millikan's account.“ (Stegmann 2009, 868)

Die *vierte Prämisse* ergibt sich aus den beiden vorangegangenen Prämissen und bedarf keiner weiteren Erläuterung.

Diskussion von Stegmanns Argument

Das Argument wurde so rekonstruiert, dass der Schluss aus den Prämissen folgt. Es bleibt also kritisch zu prüfen, ob auch die Prämissen wahr sind.

Die *ersten beiden* Prämissen scheinen mir unbedenklich. Die erste Prämisse ist fast trivialerweise wahr und die zweite Prämisse ist insofern wahr als wahrscheinlich eine Mehrzahl von Ethologen die Lichtmuster von *Photuris versicolor* als Tiersignale, die einen Gehalt haben, betrachten würden.

Die *dritte Prämisse* ist hingegen problematisch, weil die beiden impliziten Annahmen nicht zutreffend sind. Es ist zunächst plausibel anzunehmen, dass Tiersignale in der Reichweite von

²¹³ Ich habe in den Zitaten "sentence" durch "intentional icon" ersetzt. Siehe dazu auch Bemerkung 1 in Abschnitt 2.1.1.2.

Millikans Theorie liegen sollten²¹⁴. Die Annahme scheint allein schon durch die vielen Beispiele (Bienentanz, Kellenschlag des Biebers etc.), die Millikan zur Illustration ihrer Theorie verwendet, berechtigt zu sein. Da Millikan zudem innerhalb ihrer Allgemeinen Zeichentheorie (General Theory of Signs), die als grundlegend zu betrachtende Definition von *intentional icon* gibt, ist es naheliegend anzunehmen, dass Tiersignale nach Millikan *intentional icons* sein müssen. Ebenso könnte man annehmen, dass Millikan mit der Definition von *intentional icon* zugleich angibt, was für sie überhaupt als Gehaltsträger in Frage kommt, so dass etwas, das dieser Definition nicht entspricht, keinen Gehalt hat. Beide Annahmen sind aber, wie gesagt, nicht zutreffend.

Millikan weist gleich zu Beginn ihrer Allgemeinen Zeichentheorie daraufhin, dass es ihrer Ansicht nach keine allgemeine Definition eines Zeichens geben kann. D.h. es gibt keine Liste von Merkmalen, die allem und nur dem zukommt, was wir als Zeichen bezeichnen (Vgl. Millikan 1984, 85-86). Stattdessen wird nach Millikan der Zeichenbegriff durch ein Paradigma zusammengehalten. Das Paradigma bilden artikulierte²¹⁵, konventionelle Zeichen. Alle anderen Zeichen sind „Zeichen“, weil sie ein oder mehrere wichtige Merkmale mit dem Paradigma teilen.

„[...] the notion ‚sign‘ is unified not by a definition but by a paradigm.“ (Millikan 1984, 86)

„[...] articulate conventional signs, indicative and imperative, are the paradigm cases of signs. All other signs are ‚signs‘ by virtue of one or another striking resemblance to these paradigms.“ (Millikan 1984, 85)

Die Definition von *intentional icon* ist also grundlegend in dem Sinn, dass sie Millikans Paradigma eines Zeichens definiert.

„There are a great many things in the world that are intentional icons. And all of the things that we are tempted to call ‚signs‘ exhibit very striking analogies to intentional icons.“ (Millikan 1984, 96)

Daraus folgt zunächst, dass nach Millikan nicht jedes Tiersignal unter die Definition von *intentional icon* fallen muss. Vor allem folgt aber, wenn man plausiblerweise annimmt, dass alle Zeichen qua Zeichen Gehaltsträger sind, dass nach Millikan nicht nur dasjenige ein Gehaltsträger, d.h. von repräsentationaler Art (representational kind), sein kann, was unter die Definition von *intentional icon* fällt. Genau dies scheint Stegmann aber vorauszusetzen, wenn er schlussfolgert, dass die Lichtmuster von *Photuris versicolor* nach Millikans Theorie keinen Gehalt haben, weil sie nicht der Definition von *intentional icon* genügen.²¹⁶

Was man allerdings nach wie vor einwenden könnte, ist, dass sich Millikans Gehaltsbestimmung auf *intentional icons* bezieht und damit nach Millikan der Gehalt der Lichtmuster von *Photuris versicolor* unbestimmt bleibt.

Könnte der Gehalt der Lichtmuster von *Photuris versicolor* nicht einfach durch den Gehalt der Lichtmuster, die nachgeahmt werden, bestimmt sein? Für die nachgeahmten Lichtmuster liefert Millikans Theorie ja einen Gehalt. Auch Stegmann spekuliert über diese Möglichkeit,

²¹⁴ Obwohl auch daraus nicht zwingend folgt, dass sie *alle* Arten von Tiersignalen erfassen können muss.

²¹⁵ gegliederte.

²¹⁶ In einer persönlichen Emailkorrespondenz hat Ulrich Stegmann mir bestätigt, dass er tatsächlich diese stillschweigende Voraussetzung macht. Meine Kritik betrachtet Stegmann allerdings nur als einen technischen Einwand, der für sich genommen nicht ganz befriedigend ist. Sein entscheidender Punkt dafür ist der Folgende: Wenn nicht-paradigmatische Beispiele auch Gehaltsträger sein können, wieviel Ähnlichkeit mit dem Paradigma muss es geben, und in welcher Hinsicht, damit ein Träger einen Gehalt hat? Ausserdem müsste man aufpassen, so Stegmann, dass diese Antwort nicht zu einer ad hoc Ausflucht für mögliche Gegen-Beispiele für Millikans Theorie wird. In beiden Punkten stimme ich Stegmann zu, Millikans Theorie ist ergänzungsbedürftig, aber m.E. gefährdet dies Millikans Lösung des Problems der Missrepräsentation nicht grundsätzlich.

weil die Evolution eines täuschenden Signalsystem plausiblerweise von der vorherigen Evolution eines kooperativen Signalsystems abhängt. Er verwirft die Möglichkeit aber wiederum mit der Begründung, dass die Lichtmuster von *Photuris versicolor* um einen Gehalt haben zu können von repräsentationaler Art (representational kind) sein müssten, was sie seiner Ansicht nach jedoch nicht sind.

„Deceiving signaling systems may well depend for their evolution on a cooperative signaling system. But in order to misrepresent, the deceiving signals still need to be of a representational kind. And in order to be of a representational kind, the respective producing and consuming devices must cooperate.“ (Stegmann 2009, 869)

Da Stegmanns Begründung nicht zutreffend ist, wie wir zuvor gesehen haben, bleibt die Möglichkeit bestehen, den Gehalt der nachahmenden Lichtmuster durch den Gehalt der nachgeahmten Lichtmuster zu bestimmen.

Bei Millikan findet sich meines Wissens keine Diskussion von Interspezies-Täuschungen durch Nachahmung von Tiersignalen. Millikan äußert sich aber zum manipulativen Gebrauch der menschlichen Sprache. Lügen haben dabei einen Gehalt, anders wäre es auch unverständlich, wie mittels einer Lüge ein Hörer missinformiert werden könnte. Millikan zufolge haben sprachliche Ausdrücke direkte Eigenfunktionen und diese sind unabhängig von den Absichten des Sprechers. D.h. der Gehalt eines sprachlichen Ausdrucks ist unabhängig vom einzelnen Gebrauch, sei dieser nun kooperativ oder manipulativ. Genau deswegen, so Millikan, kann der Sprecher den sprachlichen Ausdruck, dessen Gehalt von seiner direkten Eigenfunktion abhängt, auch zum Nachteil des Hörers zweckentfremden.

„The intention or purpose with which a speaker utters a public language device token does afford that token a derived proper function. But this derived proper function is above and beyond its direct or stabilizing function and may or may not be the same as the direct function. Thus it is that a speaker can use a language device, just as he can use a tool or one of his body parts, to perform a function or serve a purpose that is not, however, the language device's own function.“ (Millikan 1984, 52)

Die direkte Eigenfunktion des sprachlichen Ausdrucks ist aber nicht völlig unabhängig von Sprecher- bzw. Hörerzwecken im Allgemeinen. Im Gegenteil sie hängt davon ab, dass eine kritische Masse von Fällen existiert, in denen sowohl die Sprecher- als auch die Hörerzwecke erfüllt wurden.

„But although the stabilizing function of a language device is independent of the purpose of the particular speaker who utters it, it is not independent of speaker purposes in general. The survival of a public language element without change of function must depend upon there being a critical mass of occasions upon which speakers and hearers use the element such that it performs its stabilizing function in accordance with the speaker's and the hearer's purposes.“ (Millikan 1984, 52-53)

Wird die kritische Masse nicht mehr erreicht, dann hören die Menschen auf miteinander zu reden, denn durch die überhand nehmenden Manipulationen verlieren die Ausdrücke ihre Funktion und damit ihre Bedeutung. Der manipulative Gebrauch ist also abhängig vom kooperativen Gebrauch, während das Umgekehrte nicht gilt. Dies räumt dem kooperativen Gebrauch eine gewisse Priorität ein.

Lassen sich nun Millikans Überlegungen zum manipulativen Gebrauch der menschlichen Sprache auch auf Fälle von Interspezies-Täuschungen durch Nachahmung von Tiersignalen übertragen?

Nicht ohne Weiteres. Das Problem ist, dass nach Millikan sprachliche Ausdrücke reproduktiv etablierte Familien erster Ordnung (Vgl. Abschnitt 2.1.1.3.2) bilden²¹⁷. D.h. sprachliche Ausdrücke sind Reproduktionen voneinander. Genau dies trifft aber auf die Lichtmuster von *Photinus macdermotti* und *Photuris versicolor* nicht zu, denn sie werden höchstwahrscheinlich nicht erlernt.²¹⁸ Zu diesem Ergebnis kommt auch Stegmann (Vgl. Stegmann 2009, 869), der daran anschliessend feststellt, dass es Wege geben könnte den Gehalt der manipulativen Signale durch den Gehalt der kooperativen Signale zu bestimmen, z.B. indem man Millikans Begriff der Reproduktion überarbeitet. Stegmann scheint aber eine solche Überarbeitung für zwecklos zu halten, weil sie auf der Annahme basiert, dass Kooperation für die Übertragung bzw. Bestimmung des Gehalts wesentlich ist, wofür aber kein Grund vorläge.

„But such moves would be premised on the assumption that cooperation is somehow essential for conferring and determining content. No argument has been given to support this assumption.“ (Stegmann 2009, 870)

Die letzte Behauptung, dass kein Grund für die Annahme vorläge, scheint mir ungerechtfertigt. Wie Millikans Überlegungen zu sprachlichen Ausdrücken schon angedeutet haben, kommt dem ehrlichen, kooperativen Gebrauch gegenüber dem manipulativen, nicht kooperativen Gebrauch eine Priorität zu. Es ist zwar aus evolutionsbiologischer Sicht nicht so, dass universelles Lügen (Alle lügen immer) unmöglich ist, wie dies Millikans Überlegung nahezulegen scheint. Beides ist möglich, universelles Lügen und universelles Wahrsagen (Alle sagen immer die Wahrheit) aber beide Zustände sind instabil. Zu diesem Ergebnis kommt zumindest Elliott Sober (1994, 80). Aber auch für Sober besteht eine Asymmetrie zwischen Täuschungen und kooperativen Gebrauch: Täuschungen können zeitlich nicht vor dem kooperativen Gebrauch von Signalen entstehen.

„For mimicry to evolve, there must first be a model“ (Sober 1994, 80-81)

Insbesondere erhält auch nach Sober das täuschende Signal seine Bedeutung durch das nachgeahmte Signal, dass seine Bedeutung in einem kooperativen Kontext erhalten hat (Vgl. Sober 1994, 74 und 81). In diesem Sinne ist Kooperation für die Übertragung bzw. Bestimmung des Gehalts wesentlich. Zudem haben wir in Abschnitt 2.1.1.3.7 gesehen, dass die Zuordnung von einem Signal zu einer Situation innerhalb eines kooperativen Settings erklärt werden kann. In einem alternativen, nicht-kooperativen Setting ist eine solche Erklärung nicht möglich. D.h. die Auszahlungen von S. 92, Abbildung 37 ändern sich. Die Vorrichtung B (der Empfänger) erhält eine Auszahlung, wenn Situation und Effekt passen (S1-E1 und S2-E2), die Vorrichtung A (der Sender) hingegen erhält eine Auszahlung, wenn Situation und Effekt nicht passen (S1-E2 und S2-E1).

²¹⁷ Etwas vereinfacht könnte man sagen, dass Worte über Generationen hinweg erhalten bleiben, weil Kinder sie von ihren Eltern nachahmen (kopieren). Vgl. hierzu auch Millikan (1984, 23).

²¹⁸ Dasselbe gilt aber nicht unbedingt für Intraspezies-Täuschungen. Diese könnten vermutlich analog Millikans Überlegungen zum manipulativen Gebrauch der menschlichen Sprache behandelt werden. Dadurch verlöre Stegmanns Einwand an Bedeutung. In einer Ausweitung seines ursprünglichen Arguments hatte Stegmann nämlich behauptet, dass Millikans Theorie nicht nur bestimmte Mimikry-Fälle nicht erfassen kann, sondern sämtliche Fälle von Täuschungen: “Parent-offspring relations, mate choice, resource defense, and predator-prey contexts are four broad fields of animal life where selection pressures may give rise to ‘deceiving’ signaling strategies (e.g., Krebs and Davies 1993; McGregor 2005). Signals produced by a deceiving mechanism will never be consumed by devices whose normal condition is that the producing device achieves its function. Therefore, they all lack representational content on Millikan’s account.” (Stegmann 2009, 869).

		Vorrichtung A	
		A I	A II
Vorrichtung B	B I	0,1	1,0
	B II	1,0	0,1

Nun gibt es aber keine Gleichgewichte mehr. Folglich kann sich auch keine der Zuordnungen stabilisieren und zum Standard werden. Die Signale in einem solchen Signaling-Game haben dementsprechend überhaupt keinen Gehalt (Vgl. Skyrms 2010, 77-78).

„In certain cases of diametrically opposed interests it is imposible, as Kant says, for everyone to practice deception, at least in equilibrium. That is because, in equilibrium, there is no information at all in the signals.“ (Skyrms 2010, 80)²¹⁹

Im Gegensatz zu Stegmanns Behauptung ist die Kooperation also wesentlich für die Übertragung bzw. Bestimmung des Gehalts eines Signals und dies kann auch gezeigt werden. Daraus folgt zwar nicht zwingend, dass der Gehalt des parasitären Gebrauchs aus dem kooperativen abgeleitet werden muss, aber dies ist zumindest naheliegend.

Fazit:

Stegmanns Einwand kann nicht überzeugen, denn die zentrale Prämisse seiner Argumentation, dass den Lichtmustern von *Photuris versicolor* nach Millikans Theorie kein Gehalt zugeordnet werden kann, weil sie keine *intentional icons* sind, beruht auf zwei falschen Annahmen. Die erste falsche Annahme ist, dass die Lichtmuster von *Photuris versicolor*, sofern sie in Millikans Theorie erfasst werden können, unter Millikans Definition von *intentional icons* fallen müssen. Die zweite falsche Annahme ist, dass nach Millikans Theorie nur dann etwas einen repräsentationalen Gehalt haben kann, wenn es mindestens Millikans Definition von *intentional icon* genügt.

²¹⁹ Auch Habermas argumentiert dafür, dass der verständigungsorientierte Sprachgebrauch der „Originalmodus“ ist und dass sich der zweckorientierte, manipulative Sprachgebrauch diesem gegenüber parasitär verhält. „Verständigung wohnt als Telos der Sprache inne“ (Habermas 1981, 387).

3.2.8 Nanay: The individuation of trait types

In diesem Abschnitt geht es darum, einen Einwand von Bence Nanay gegen ätiologische Funktionstheorien, wie diejenige Millikans, zu rekonstruieren und anschließend zu diskutieren. Nanays Einwand ist, dass ätiologische Funktionstheorien fälschlicherweise annehmen, dass eine ihrer Voraussetzungen, nämlich die Individuation von Merkmalstypen, unproblematisch sei (Vgl. Nanay 2010). Nanay zufolge steht den ätiologischen Funktionstheorien aber keine Möglichkeit einer Individuation von Merkmalstypen zur Verfügung, die nicht entweder zirkulär ist oder die Möglichkeit von Fehlfunktionen durch missgebildete Funktionsträger ausschließt. Nanays Einwand steht somit in unmittelbarem Zusammenhang mit Millikans Lösung des Problems der Missrepräsentation, weil der Einwand Millikans Funktionsauffassung, die eine Voraussetzung ihrer Lösungsansätze ist, angreift²²⁰. Beide Alternativen sind für Millikan negativ: Der Zirkularitätsvorwurf untergräbt die naturalistischen Ansprüche ihrer Theorie und die Unmöglichkeit von Fehlfunktionen durch missgebildete Funktionsträger blockiert ihren zweiten Lösungsansatz.

Rekonstruktion von Nanays Einwand²²¹

Der Kern von Nanays Kritik auf Millikans bezogen lässt sich meines Erachtens wie folgt rekonstruieren:

1. Eine Funktionstheorie, die zirkulär ist oder Fehlfunktionen durch missgebildete Funktionsträger ausschließt, ist fehlerhaft.
2. Ätiologische Funktionstheorien setzen die Individuation von Merkmalstypen voraus.
3. Die Möglichkeiten der Individuation von Merkmalstypen, die ätiologischen Funktionstheorien zur Verfügung stehen, sind entweder zirkulär oder schliessen die Möglichkeit von Fehlfunktionen durch missgebildete Funktionsträger aus.
4. Ätiologische Funktionstheorien sind fehlerhaft.

Der Bezug zu Millikan ergibt sich durch das weitere Argument:

5. Millikans Lösungsansätze basieren auf einer ätiologischen Funktionstheorie.
6. Millikans Lösungsansätze basieren auf einer fehlerhaften Funktionstheorie (und sind deshalb selbst fehlerhaft).

Nachdem wir die Struktur des Argumentes rekonstruiert haben, wollen wir nun die einzelnen Prämissen genauer anschauen, wobei wir uns auf die ersten drei aus Nanays eigentlichem Argument beschränken.

Auf die *erste Prämisse* soll nicht weiter eingegangen werden, da sie an sich plausibel ist und auch Millikans eigenem Anspruch an ihre Funktionstheorie entspricht.

²²⁰ In einem kürzlich erschienen Aufsatz hat Nanay seine Kritik auch explizit in diesen Kontext gesetzt (Vgl. Nanay 2014). Ansonsten wiederholt Nanay (2014) aber lediglich seine alte Kritik, teilweise sogar wortwörtlich.

²²¹ Dieser Teil basiert auf einem Vortrag, den ich zusammen mit Christian Steiner in Fribourg 2013 gehalten habe.

Die *zweite Prämisse* möchte ich hingegen kurz erläutern. Ätiologische Funktionstheorien setzen die Individuation von Merkmalstypen insofern voraus, als deren Definition von Funktion auf bestimmte Effekte referiert, die von vergangenen tokens *desselben Merkmalstyps* hervorgerufen wurden (Vgl. Nanay 2010, 413).

„The trait whose function is to be defined and the traits that have been selected for in the past must be of the same type.“ (Nanay 2010, 415)

In Millikans Theorie und Terminologie heisst das, dass das Haben einer Eigenfunktion die Zugehörigkeit zu reproduktiv etablierten Familie voraussetzt und dies entspricht voll und ganz unserer Darstellung (Vgl. Abschnitt 2.1.1.3.3).

Die *dritte Prämisse* macht zweifelsohne den Kern von Nanays Argument aus und soll deshalb ausführlicher behandelt werden. Nanay sieht drei Möglichkeiten der Individuation von Merkmalstypen (*trait type individuation*), zwei führen zur Zirkularität und die dritte erlaubt keine Fehlfunktionen durch missgebildete Funktionsträger:

- i) Kriterium: Funktion (zirkulär)
- ii) Kriterium: Morphologie (schliesst missgebildete Funktionsträger aus)
- iii) Kriterium: Homologie (zirkulär)

Dass das *erste Kriterium* in eine Zirkularität führt, ist offensichtlich. Wenn wir die Individuation von Merkmalstypen durch deren Funktion vornehmen (Herzen sind Dinge, die die Funktion haben Blut zu pumpen), und dann die so individuierten Merkmalstypen für die Definition von Funktion benutzen, dann produzieren wir einen Zirkel.

„When we are explaining function, the claim that x^* (the trait whose function we are explaining) is a token of type X (the traits that have been selected in the past) is part of the explanans. Hence, we cannot use the explanandum (function) to explain part of the explanans (why x^* is a token of type X).“ (Nanay 2010, 416)

Das *zweite Kriterium* schliesst missgebildete Funktionsträger aus. Es ist ja gerade die Eigenschaft eines missgebildeten Herzens, nicht alle wesentlichen morphologischen Eigenschaften eines funktionstüchtigen Herzen zu haben. D.h. die Menge der Herzen muss wegen der Bedingung, dass Fehlfunktionen durch missgebildete Funktionsträger möglich sein müssen, morphologisch inhomogen sein, weswegen ein morphologisches Kriterium hier nicht zielführend sein kann.

Nach dem *dritten Kriterium* sind zwei Merkmalstokens vom selben Merkmalstyp, wenn sie homologe Merkmale sind. „Ein Merkmal zweier oder mehrerer Taxa ist homolog, wenn es sich von demselben (oder einem entsprechenden) Merkmal ihres nächsten gemeinsamen Vorfahren ableitet.“ (Mayr 1998, 189)²²². Das Problem mit diesem Kriterium ist, dass der Flügel eines Adlers und der Vorderfuss eines ausgestorbenen Lurches homologe Merkmale sind, woraus folgen müsste, dass sie dieselbe Funktion haben. Genau dies ist nach Nanay aber mehr als abwegig, denn wir müssten dann dem Adlerflügel die Funktion zuschreiben eine kriechende Fortbewegung zu unterstützen.

„More importantly, such broad way of typing traits does not help in our definition of function as using this way of talking about trait types would attribute the function of crawling to the eagle's wings.“ (Nanay 2010, 418)

Um diesen Einwand zu entkommen, könnte man dafür plädieren, dass nur homologe Merkmale der jüngeren Vergangenheit zählen. Zwei Merkmale wären dann vom selben Typ, wenn sie homologe Merkmale der jüngeren Vergangenheit (recent homologues) sind (Vgl. Nanay 2010, 418). Aber auch dieser Ausweg bringt uns nicht weiter, denn es gibt keine

²²² “homology—similarity due to shared ancestry or common descent” (Neander und Rosenberg 2012, 2).

Möglichkeit „recent“ in diesem Zusammenhang sinnvoll zu definieren ohne darauf Bezug zu nehmen, wofür die Merkmale selektiert wurden, d.h. auf ihre Funktion.

„The problem with this suggestion is that there is no non-circular way of cashing out what is meant by 'recent' here. [...] what 'recent' amounts to depends on what the trait in question has been selected for; what 'recent' amounts to depends on the etiological function of the trait.“ (Nanay 2010, 418)

Mit anderen Worten: Das dritte Kriterium entspricht dem ersten Kriterium und dieses ist zirkulär (Vgl. Nanay 2010, 419).

Diskussion von Nanays Argument

Das Argument wurde so rekonstruiert, dass der Schluss aus den Prämissen folgt. Es bleibt also kritisch zu prüfen, ob auch die Prämissen wahr sind. Ich konzentriere mich hierbei auf die dritte Prämisse, die auch im Zentrum von Neanders und Rosenbergs Erwiderung auf Nanays Einwand steht (Neander & Rosenberg 2012), die ich zunächst wiedergeben möchte. Nanay hat seinerseits auf diese Erwiderung reagiert (Nanay 2012) und diese Reaktion führt uns direkt zu einer Schwierigkeit, von der auch Millikans Theorie der Eigenfunktion unmittelbar betroffen ist. Nanay selbst bezieht sich kaum direkt auf Millikans Theorie. Der Grund hierfür ist wahrscheinlich seine Ablehnung von gen-zentristischen Auffassungen zu homologen Merkmalen.

„To make this suggestion as plausible as possible, we should not confuse it with the view that homologous traits are ‘coded by’ the same gene, for the simple reason that no trait should be taken to be ‘coded by’ a gene. The way a trait turns out depends partly on the gene, but it also depends on the intra- and extra-cellular environment during the developmental process. Assuming that the genetic factor can be singled out from this complex causal network is biologically very implausible.“ (Nanay 2010, 417)

Die abgelehnte Sicht entspricht in etwa derjenigen von Millikan und den Kritikpunkt von Nanay an dieser Sicht haben wir bereits diskutiert (Vgl. Abschnitt 2.1.1.3.5 und 3.2.4).

Betrachten wir also zunächst Neander und Rosenbergs Erwiderung. Ihr wesentlicher Kritikpunkt an Nanays Einwand ist, dass Nanay fälschlicherweise voraussetzt, dass eine Berufung auf das, wofür etwas selektiert wurde (selected for), immer bereits eine Berufung auf eine Funktion im ätiologischen Sinne ist (Vgl. Neander und Rosenberg 2012, 620, Fn 9). Korrekt ist hingegen nach Neander und Rosenberg, dass die Merkmalsindividuation und die Funktionszuschreibung zugleich ausgeführt werden.

“Traits are typed *in* ascribing function to them, not before ascribing function to them.”
(Neander und Rosenberg 2012, 620)

Beide, die Merkmalsindividuation und die Funktionszuschreibung, ko-supervenieren nach Neander und Rosenberg auf zwei zugrunde liegenden Fakten:

- a) Abstammungslinien und
- b) Selektionsdrücke, die auf diese Linien eingewirkt haben.

Neander und Rosenberg versuchen nun Nanays Zirkularitätsproblem in vier Schritten zu lösen (Vgl. Neander und Rosenberg 2012, 618-621).

Der *erste Schritt* besteht darin, das betreffende Merkmalstoken in einer Abstammungslinie zu lokalisieren. Ein Beispiel mag hier hilfreich sein und um die Sache möglichst einfach zu

halten, nehmen wir an, dass das betreffende Merkmalstoken nur einen Vorfahren hat²²³. Die Flosse eines Pinguins (das betreffende Merkmalstoken) ist dann mit der Flosse (Merkmalstoken) seines Vorfahren über die Vererbungsmechanismen verbunden. Die Flosse des Vorfahren ist wiederum mit der Flossen von dessen Vorfahren verbunden usw. Die Linie der verbundenen Merkmalstokens geht dann weit zurück in die Vergangenheit, wie in Abbildung 73 angedeutet.

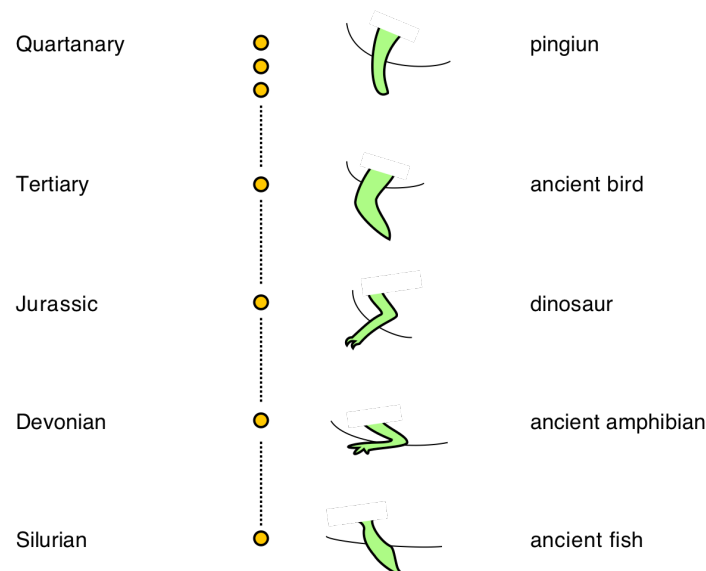


Abbildung 73: Abstammungslinie einer Pinguinflosse

Der *zweite Schritt* betrifft die Selektionsdrücke, die auf die Linie wirken bzw. wirkten. Jedes Glied in der Abstammungslinie war mit verschiedenen Umweltbedingungen konfrontiert. Diese Bedingungen änderten sich über die Zeit und erzeugten so verschiedene Selektionsdrücke auf die Glieder der Linie, was über den Prozess der Natürlichen Selektion zu Adaptionen, wie Flossen, Beinen und Flügeln, führte (siehe Abbildung 74).

²²³ D.h. wir gehen von einer monoparentalen Vererbung aus. Vgl. aber Abschnitt 3.2.4 zu den Problemen, die Millikans Theorie mit der sexuellen Vermehrung hat.

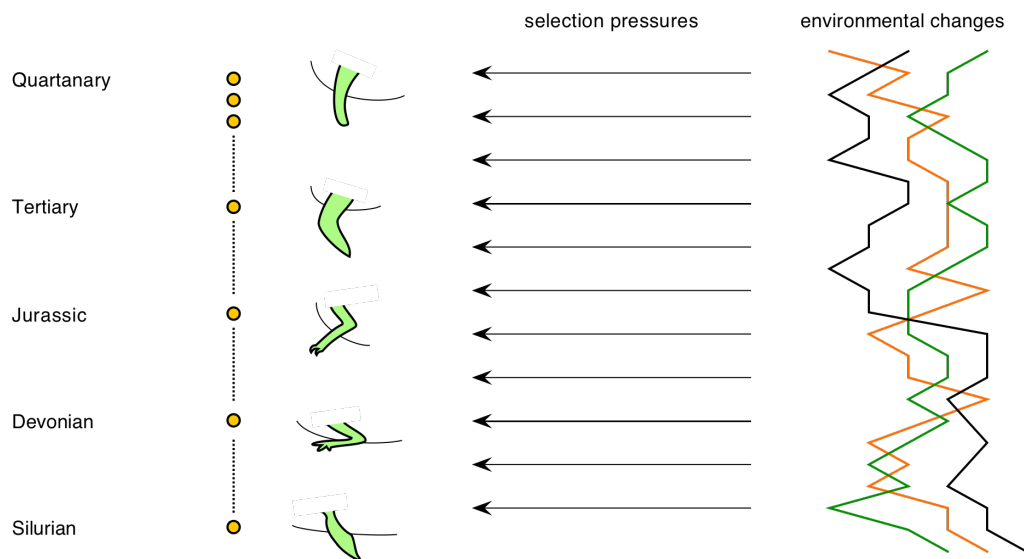


Abbildung 74: Abstammungslinie einer Pinguinflosse mit sich verändernden Selektionsdrücken, die auf die Linie eingewirkt haben

Der *dritte Schritt* besteht darin, die Abstammungslinie entsprechend bestimmter Selektionsdrücke in Segmente zu unterteilen. D.h. wir denken uns Abtrennungslinien an den Stellen, an denen der Selektionsdruck für eine schwimmende Fortbewegung beginnt und wo er endet (analog für laufende und fliegende Fortbewegung). Wir erhalten dann segmentierte Abstammungslinien wie sie in Abbildung 75 exemplarisch dargestellt sind.

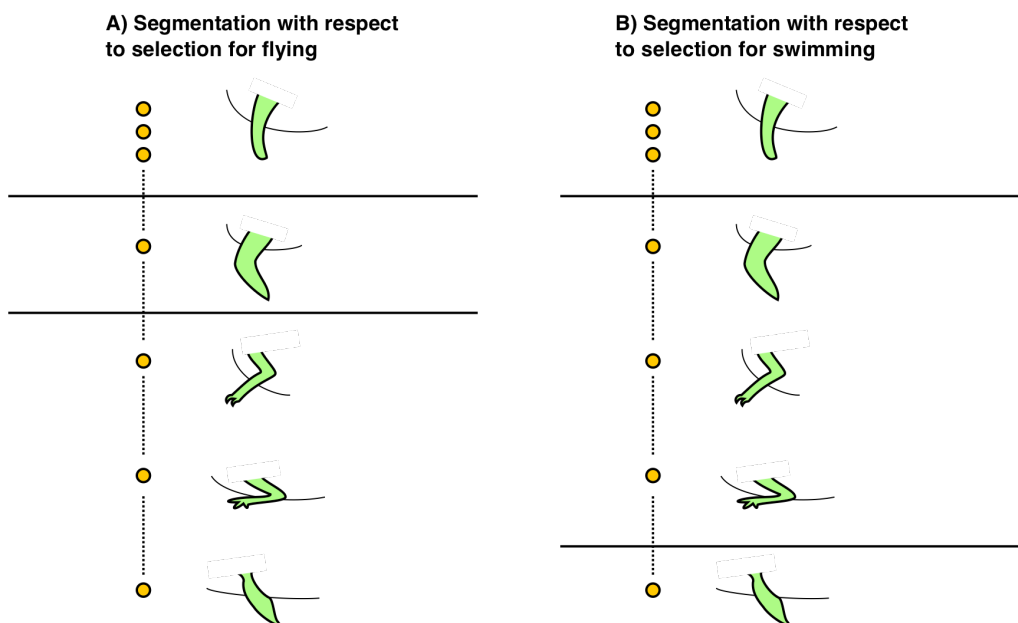


Abbildung 75: Segmentierte Abstammungslinien. A) Segmentierung nach einem Selektionsdruck für fliegende Fortbewegung. B) Segmentierung nach einem Selektionsdruck für schwimmende Fortbewegung

Der *vierte und letzte Schritt* zeigt wie die Merkmalsindividuation und die Funktionszuschreibung zugleich vorgenommen werden. Wenn wir nämlich nun ein

Merkmalstoken in einem Segment der Abstammungslinie herausgreifen, sagen wir *zugleich* zweierlei, erstens „Dies ist ein Flügel“ und zweitens „Dies hat die Funktion Fliegen zu ermöglichen“. Nanay liegt also falsch, wenn er annimmt, dass die Merkmalsindividuation die Funktionszuschreibung voraussetzen würde, sie findet einfach gleichzeitig statt.

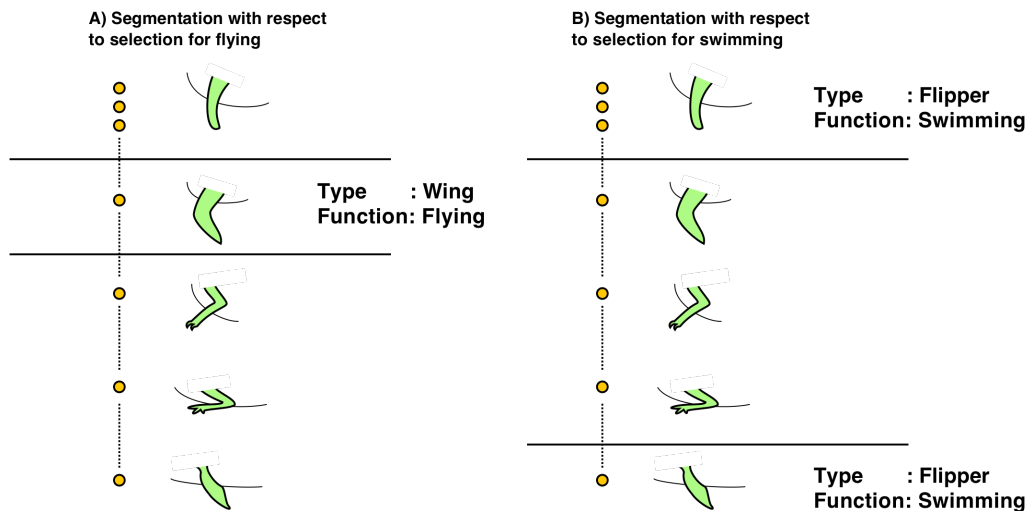


Abbildung 76: Gleichzeitige Merkmalsindividuation und Funktionszuschreibung

Nanay hat auf diesen Lösungsvorschlag reagiert. Das Problem mit diesem Vorschlag ist Nanay zufolge, dass die Abstammungslinien (aus Schritt 1) nicht identifiziert werden können, ohne die Merkmalstypen bereits vorauszusetzen. Das Zirkularitätsproblem bleibt somit bestehen²²⁴.

„The problem with this argument is that lineages themselves could not be identified without talking about trait types. [...] the circularity worry has not gone away.“ (Nanay 2012, 624)

Neander und Rosenberg hatten gesagt, dass die Glieder der Abstammungslinie über die Vererbungsmechanismen verbunden sind.

„How are token traits located in lineages? The mechanisms and processes that are responsible for inheritance link traits in ancestral and descendant relations [...].“ (Neander and Rosenberg 2012, 618)

Wie sollen wir diese Prozesse und Mechanismen der Vererbung verstehen? Funktional? Dann droht wieder das Zirkularitätsproblem. Nicht-funktional? Dann droht wieder der Ausschluss von missgebildeten Funktionsträgern. Wie verhält sich dies in Millikans Theorie?

Millikans versteht diese Prozesse zunächst nicht-funktional. Die Entstehung von Reproduktionen und auch von reproduktiv etablierten Familien erster Ordnung ist bei Millikan unabhängig von irgendwelchen Funktionen denkbar. Wir wollen jetzt genauer untersuchen, inwiefern dies zu Schwierigkeiten bezüglich der Möglichkeit von missgebildeten Mitgliedern von reproduktiv etablierten Familien führt.

Im Abschnitt 2.2.2.2 hatten wir bereits festgestellt, dass sich Millikans Aussagen zu missgebildeten REF-Mitgliedern *nur* auf HOREF-Mitglieder zu beziehen. Millikan hatte ihre Definition der HOREF durch eine dritte Version ergänzt, die explizit Raum für die

²²⁴ Es entsteht ohnehin etwas der Eindruck, dass Neander und Rosenberg hier eine *bootstrapping strategy* verfolgen. In der These von der Gleichursprünglichkeit der Merkmalsindividuation und der Funktionszuschreibung erscheint die alte Zirkularität in neuem Gewand.

Möglichkeit missgebildeter REF-Mitglieder schaffen soll. Millikans Grundidee, die Frage nach der Mitgliedschaft in einer REF durch die Herkunft, die Vergangenheit eines Objektes und nicht durch dessen tatsächliche Eigenschaften zu entscheiden, ist an sich plausibel. Aber diese Grundidee wird durch die, in die dritte Definitionsversion eingebaute Vagheit, geschwächt. Weiterhin hatten wir festgestellt, dass man hier kritisch fragen könnte, ob eine solche Bedingung nicht im vornherein immun für Gegenbeispiele ist und somit die Frage der Zugehörigkeit zu einer reproduktive etablierten Familie nicht zu einer rein subjektiven Angelegenheit wird. Zudem betreffen die missgebildeten HOREF-Mitglieder vom Kontext her die Ontogenese, wo uns oft das genaue Wissen, sprich die Normale Erklärung fehlt (Vgl. Abschnitt 2.1.1.3.5 und 3.2.4). Alles in allen – so war unser Fazit – ist einige Skepsis gegenüber Millikans Bestimmung von missgebildeten HOREF-Mitgliedern angebracht (Vgl. 118-120).

Wir wollen uns aber vorerst eines abschließenden Urteils dazu enthalten und zunächst fragen, ob es auch missgebildete FOREF-Mitglieder geben kann. Sowohl HOREF- als auch FOREF-Mitglieder können Eigenfunktionen haben und die Eigenfunktionen von FOREF-Mitgliedern sind sogar die grundlegenderen Eigenfunktionen²²⁵. Wenn also Eigenfunktionen keine Dispositionen sind, wie Millikan immer wieder betont hat (Vgl. z.B. Millikan 1984, 93 und 1993b, 32), dann sollte es auch missgebildeten FOREF-Mitglieder geben, andernfalls wären Eigenfunktionen von FOREF-Mitgliedern nicht von Dispositionen zu unterscheiden. Das Zustandekommen von missgebildeten FOREF-Mitgliedern zu erläutern, erweist sich jedoch als schwierig²²⁶. Die Mitgliedschaft in einer HOREF wurde vor allem über die Eigenfunktion einer weiteren REF bestimmt, die eben darin besteht Mitglieder der HOREF zu produzieren. Ein leicht abweichender Produktionsprozess könnte dann zu einem missgebildeten HOREF-Mitglied führen (Vgl. Abbildung 54, S. 119). Ein ähnliches Szenario funktioniert aber bezüglich der FOREFs nicht. Angenommen bei dem Prozess, bei dem unter der Beteiligung einer RNA 1 eine weitere RNA 2 hervorgeht, hat die RNA 2 nicht diejenige Eigenschaft C²²⁷ relativ zu der die FOREF, zu der die RNA 1 gehört, bestimmt ist, dann ist die RNA 2 kein missgebildetes Mitglied der FOREF, sondern überhaupt kein Mitglied der FOREF (siehe Abbildung 78).

²²⁵ Ohne Eigenfunktionen von FOREFs kann es keine Eigenfunktionen von HOREFs geben, was umgekehrt nicht gilt.

²²⁶ Auf diese Schwierigkeit wurde ich erstmalig durch Christian Steiner aufmerksam gemacht. Vgl. hierzu auch Christian Steiner „Troubles for Etiological Accounts of Normative Function“ (in Vorbereitung).

²²⁷ Die Eigenschaft C resultiert aus der Basensequenz der RNA, z.B. die Eigenschaft für ein bestimmtes Protein zu codieren. Die Eigenschaft C bewirkt dann unter Normalen Bedingungen den Effekt E, z.B. die Synthese des Proteins.

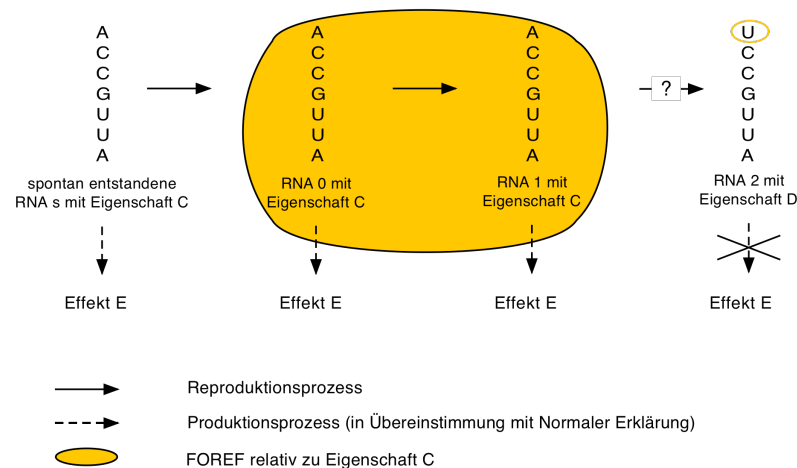


Abbildung 77: RNA 2 ist keine Reproduktion (C) der RNA 1 und damit kein Mitglied der FOREF (C).

Millikan kann bezüglich der FOREFs nicht auf so etwas wie einen Reproduktionsfehler zurückgreifen. Ein Prozess der nicht die relevanten Eigenschaften von der RNA 1 auf die RNA 2 überträgt ist überhaupt kein Reproduktionsprozess auch kein fehlerhafter. Um von einem fehlerhaften Reproduktionsprozess sprechen zu können, müsste dieser als funktional aufgefasst werden. Eine solche Auffassung wäre jedoch für Millikan fatal, denn dann würde der Begriff der Reproduktion schon den Begriff Eigenfunktion²²⁸ voraussetzen und eine Definition von Eigenfunktion mit Hilfe des Begriffes der Reproduktion wäre zirkulär. Genau dies ist Nanays Einwand im Kontext von Millikans Theorie. Missgebildete FOREF-Mitglieder erweisen sich also für Millikans Theorie als schwierig.

Meines Erachtens gibt es aber eine Ausweg aus dieser Schwierigkeit. Eine Möglichkeit doch von missgebildeten FOREF-Mitgliedern sprechen zu können, besteht darin, dass die betreffenden Objekte erst *nach* dem Reproduktionsprozess verändert werden. So könnte die RNA 2 die Reproduktion der RNA 1 sein, worauf sich ihre Mitgliedschaft in einer bestimmten FOREF begründet. Diese RNA 2 könnte nun zu einem *späteren* Zeitpunkt durch äußere Einflüsse (z.B. UV-Strahlung) eine Deformation erfahren²²⁹, z.B. die Bildung eines Pirimidindimers (siehe Abbildung 79), die dazu führt, dass die RNA 2 die für die FOREF relevante Eigenschaft C (Codieren für Protein P) nicht mehr besitzt und folglich die Eigenfunktion der FOREF, den Effekt E zu bewirken (die Synthese des Proteins), nicht ausüben kann, auch nicht unter den relevanten Bedingungen.

²²⁸ Oder irgendeinen anderen nicht-dispositionalen Funktionsbegriff.

²²⁹ Die UV-Strahlung bewirkt eine vermehrte Bildung freier Radikale, die wiederum zu oxidativen Schäden an RNA (Vgl. Bernard et al. 2012) und DNA (Vgl. Marnett 2000, Wiseman und Halliwell 1996) führen.

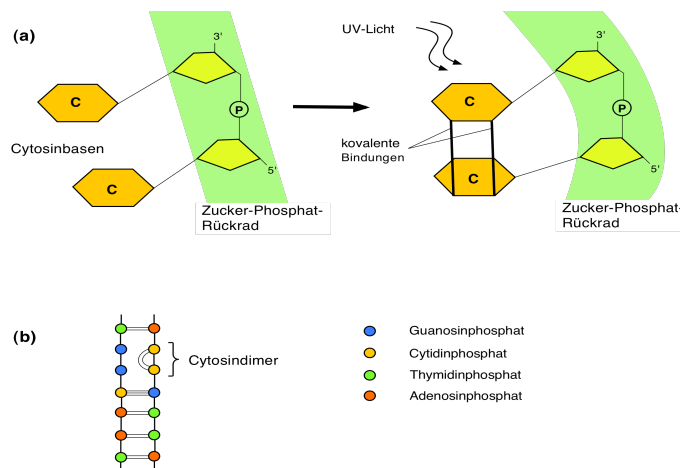


Abbildung 78: (a) Entstehung eines Pyrimidindimers, (b) deformierte DNA (In Anlehnung an Pierce, A. Benjamin. 2007. *Genetics. A Conceptual Approach*. New York: W.H. Freeman and Company. 488).

Eine solche RNA könnte als missgebildetes FOREF-Mitglied gelten, wobei eben die Herkunft wichtiger ist als die aktuellen Eigenschaften.

Auf ähnliche Weise könnten auch missgebildete HOREF-Mitglieder entstehen. So könnte sich zum Beispiel ein bestimmtes Organ während der Ontogenese vollfunktionstüchtig ausbilden und später durch eine Krankheit oder eine Verletzung die Eigenschaften verlieren, die für die Ausübung seiner Eigenfunktion notwendig sind²³⁰. Ein solches Organ würde m.E. ebenfalls zu Recht als ein missgebildetes HOREF-Mitglied gelten.

Wir können nun unsere Überlegungen zur Möglichkeit missgebildeter REF-Mitglieder in folgender Tabelle zusammenfassen:

Entstehung eines missgebildeten REF-Mitglieds möglich?	
Beim Reproduktionsvorgang (FOREF)	<i>nein</i>
Beim Produktionsvorgang (HOREF)	<i>sehr fragwürdig</i>
Nach Reproduktionsvorgang (FOREF)	<i>ja</i>
Nach Produktionsvorgang (HOREF)	<i>ja</i>

Es kann also missgebildete FOREF-Mitglieder geben, sobald deren Entstehen nach dem Reproduktionsprozess mit in Betracht gezogen wird. Dasselbe gilt für HOREF-Mitglieder, so dass ein endgültiges Urteil, ob auch beim Produktionsvorgang einer HOREF ein missgebildetes Mitglied entstehen kann, nicht entscheidend ist.

Dementsprechend gilt für die Eigenfunktionen von FOREF- und HOREF-Mitgliedern, dass sie keine Dispositionen sind, denn sowohl FOREF- als auch HOREF-Mitglieder können Eigenfunktionen haben, ohne tatsächlich in der Lage zu sein diese unter den relevanten Bedingungen auszuüben.

²³⁰ Hierfür muss man aber zwischen Entwicklung (Ontogenese) und Veränderung im Allgemeinen unterscheiden, wie etwa Mahner und Bunge (Vgl. Mahner und Bunge 200, 265f).

FAZIT:

Nanays Einwand kann nicht vollständig entkräftet werden. Die Erläuterung der Entstehung von missgebildeten oder allgemeiner abNormalen REF-Mitgliedern während des Reproduktions- oder Produktionsprozesses ist schwierig. Entweder ist sie sehr vage oder zirkulär. Die Entstehung von abNormalen REF-Mitgliedern kann aber problemlos nach dem Reproduktions- oder Produktionsprozess erläutert werden. Dieser Ausweg hat allerdings die Konsequenz, dass *Missbildungen* im eigentlichen Sinne in Millikans Theorie keine Missrepräsentation hervorrufen können.

3.2.9 Wunder: Verfehlen der Funktion unter abNormalen Bedingungen

In diesem Abschnitt geht es mir darum einen eigenen Einwand gegen Millikans Theorie zu präsentieren. Mein Einwand lautet:

Wenn dispositionale Funktionsbegriffe kein Verfehlen der Funktion zulassen, dann gibt es auch kein Verfehlen der Funktion unter abNormalen Bedingungen in Millikans Sinne.

Aus Letzterem wiederum folgt, dass Millikans Lösungsansatz für das Problem der Missrepräsentation über umweltbedingte Missrepräsentationen scheitert.

Die Voraussetzung meines Einwandes, dass dispositionale Funktionsbegriffe kein Verfehlen der Funktion zulassen, ist insofern unproblematisch als Millikan diese Auffassung selbst vehement vertreten hat.

„[...] a crippling defect of any definition that looks for function in current dispositions rather than in history is that such definitions cannot ground the notion of malfunction. [...] Especially important, no such notion of function could ground an understanding of the nature of cognitive mistakes.“ (Millikan 1993d, 32)²³¹

Ich möchte meinen Einwand in drei Schritten erläutern.

Im *ersten Schritt* lege ich dar, wieso ein dispositionaler Funktionsbegriff ein Verfehlen der Funktion durch einen Funktionsträger unmöglich macht.

Im *zweiten Schritt* wiederhole ich, wie Millikan ein Verfehlen der Funktion durch einen Funktionsträger aufgrund von abNormalen Umweltbedingungen erklärt.

Im *dritten Schritt* zeige ich, dass Millikans Erklärung des Verfehlens der Funktion aufgrund von abNormalen Umweltbedingungen mit einem dispositionalen Funktionsbegriff vollkommen kompatibel ist. Daraus folgt: Millikans Erklärung erweckt nur den Eindruck eines Verfehlens, das tatsächlich aber nicht besteht.

1 Dispositionale Funktionsbegriffe machen ein Verfehlen der Funktion unmöglich

Ziel dieses Teilabschnittes ist es darzulegen, wieso ein dispositionaler Funktionsbegriff ein Verfehlen der Funktion durch einen Funktionsträger unmöglich zu machen scheint.

Hierfür müssen wir zunächst klären, wann wir einem bestimmten Objekt überhaupt eine Disposition zuschreiben. Nehmen wir beispielsweise die Disposition wasserlöslich zu sein. Wir sagen, ein Objekt ist wasserlöslich, wenn es sich auflöst, nachdem es in Wasser gegeben wurde. Dabei wird üblicherweise davon ausgegangen, dass wir für jede Disposition bestimmte Auslösebedingungen (stimulus conditions, SC) und bestimmte Manifestationen (manifestations, M) identifizieren können (Vgl. Choi u. Fara 2012). D.h. jeder Dispositionsträger zeigt charakteristische Manifestationen (M) unter bestimmten Bedingungen (SC). Im Fall der Wasserlöslichkeit ist die Bedingung (SC), dass das Objekt in Wasser gegeben wird, und die Manifestation (M) besteht darin, dass das Objekt sich auflöst.

Für unsere weiteren Überlegungen ist es nun hilfreich alle möglichen Kombinationen der Ausprägungen von Auslösebedingungen (erfüllt/nicht erfüllt) und Manifestationen (eingetreten/nicht eingetreten) aufzulisten. Wir erhalten vier Fälle (siehe Abbildung 79).

²³¹ Vgl. auch: "For the categories 'heart,' 'liver,' 'eye,' 'brain,' and also 'idea,' 'belief,' and 'intention' are proper function categories, defined in the end by reference to long-term and short-term evolutionary history, not present constitution or disposition. Were this not so, there could not be malformed hearts or nonfunctioning hearts nor could there be confused ideas or empty ideas or false beliefs, etc." (Millikan 1984, 93)

	Auslösebedingung (SC): Objekt im Wasser	Manifestation (M): Objekt löst sich auf
Fall 1:	W	W
Fall 2:	W	F
Fall 3:	F	W
Fall 4:	F	F

W : erfüllt / eingetreten
F : nicht erfüllt / nicht eingetreten

Abbildung 79: Die vier möglichen Kombinationen der Ausprägungen von Auslösebedingung und Manifestation (Beispiel: Wasserlöslichkeit)

Gehen wir zunächst von Einzelbeobachtungen aus, in denen einer der vier Fälle eintritt. Beginnen wir mit Beobachtungen, in denen die letzten beiden Fälle eintreten. Bei diesen Beobachtungen sind also die Auslösebedingungen nicht erfüllt. Solche Beobachtungen scheinen *irrelevant* dafür zu sein, einem Objekt eine Disposition zu zuschreiben. Ob ein Objekt wasserlöslich ist oder nicht, zeigt sich letztlich erst, wenn es in Wasser gegeben wird. Vor allem erwarten wir auch von einem wasserlöslichen Objekt nicht, dass es ein Sich-Auflösen manifestiert, wenn es gar nicht in Wasser gegeben wurde. Aufgrund von Einzelbeobachtungen, in denen die letzten beide Fälle eintreten, lässt sich also nicht sagen, ob ein Objekt wasserlöslich ist oder nicht. Anders verhält es sich mit Beobachtungen, in denen die ersten beiden Fälle eintreten. Diese sind für Zuschreibung von Dispositionen relevant. Beobachten wir, dass der erste Fall eintritt, dann schreiben wir dem Objekt die betreffende Disposition zu, tritt hingegen der zweite Fall ein, sprechen wir dem Objekt die betreffende Disposition ab. Ein Objekt, das sich auflöst, nachdem es in Wasser gegeben wurde, ist wasserlöslich. Während ein Objekt, das sich nicht auflöst, nachdem es in Wasser gegeben wurde, nicht wasserlöslich sein kann.

Mit anderen Worten: Das Haben einer Disposition besteht darin, diese unter bestimmten Bedingungen zu manifestieren. Dementsprechend ist es unmöglich, dass ein Objekt zwar eine bestimmte Disposition hat, aber diese nicht unter den relevanten Bedingungen manifestiert. In diesem Sinne ist ein Auseinandertreten von Disposition-Haben und Disposition-Manifestieren unmöglich.

Nun beruhen unsere Zuschreibungen von Dispositionen selten auf Einzelbeobachtungen. Wir sollten daher überlegen, welchen Einfluss mehrfache Beobachtungen, die mehr als einen der vier verschiedenen Fälle enthalten, auf unsere Beurteilung haben.

Angenommen wir machen folgende Beobachtungen von Objekten des Typs A (Serie 1):

Beobachtungen (Serie 1) von Objekten des Typs A		Fall in Bezug auf Wasserlöslichkeit
Typ I:	Objekt löst sich auf nachdem es in Wasser gegeben wurde	Fall 1 (WW-Fall)
Typ II:	Objekt zeigt keine Veränderung nachdem es in Flüssigkeit X gegeben wurde	Fall 4 (FF-Fall)
Typ III:	Objekt zieht sich zusammen nachdem es Temperaturen über 75°C ausgesetzt wurde	Fall 4 (FF-Fall)

Abbildung 80: Serie 1 von Beobachtungen mit drei Beobachtungstypen und die Angabe welchem Fall jeder Beobachtungstyp hinsichtlich der Disposition wasserlöslich zu sein entspricht

Hinsichtlich der Frage, ob Objekte des Typs A wasserlöslich sind, entsprechen die Beobachtungen vom Typ I dem Fall 1 (siehe Abbildung 79), die Auslösebedingung ist erfüllt und die Manifestation ist eingetreten, und Beobachtungen vom Typ II und III dem Fall 4 (siehe Abbildung 79), weder die Auslösebedingung ist erfüllt und noch ist die Manifestation eingetreten. Die Beobachtungsserie 1 enthält also die Fälle 1 und 4. Aufgrund der Beobachtungen vom Typ I (Fall 1) würden wir Objekten des Typs A die Disposition wasserlöslich zu sein zweifelsohne zuschreiben. Die anderen Beobachtungen (Fall 4) sind für dieses Urteil irrelevant.

Nehmen wir nun an, wir hätten statt den eben betrachteten Beobachtungen Folgendes (Serie 2) beobachtet:

Beobachtungen (Serie 2) von Objekten des Typs A		Fall in Bezug auf Wasserlöslichkeit
Typ I:	Objekt löst sich auf nachdem es in Wasser gegeben wurde	Fall 1 (WW-Fall)
Typ II:	Objekt löst sich auf nachdem es in Flüssigkeit X gegeben wurde	Fall 3 (FW-Fall)
Typ III:	Objekt zieht sich zusammen nachdem es Temperaturen über 75°C ausgesetzt wurde	Fall 4 (FF-Fall)

Abbildung 81: Serie 2 von Beobachtungen mit drei Beobachtungstypen und die Angabe welchem Fall jeder Beobachtungstyp hinsichtlich der Disposition wasserlöslich zu sein entspricht

Die Beobachtungsserie 2 enthält also die Fälle 1, 3 und 4. Aufgrund der Beobachtungen vom Typ I (Fall 1) können wir wieder Objekten des Typs A die Disposition wasserlöslich zu sein zuschreiben. Interessant sind hier die Beobachtungen vom Typ II, die dem Fall 3 (siehe Abbildung 79) entsprechen. Sie lassen uns vermuten, dass sich Objekte des Typs A generell in bestimmten Flüssigkeiten (Flüssigkeiten des Typs Z) auflösen, zu denen auch Wasser und die Flüssigkeit X gehören. Trifft diese Vermutung zu, dann sagen wir, dass Objekte vom Typ A die Disposition haben sich in Flüssigkeiten des Typs Z aufzulösen.

Nehmen wir abschließend an, wir hätten Folgendes (Serie 3) beobachtet:

Beobachtungen (Serie 3) von Objekten des Typs A		Fall in Bezug auf Wasserlöslichkeit
Typ I:	Objekt löst sich auf nachdem es in Wasser gegeben wurde	Fall 1 (WW-Fall)
Typ II:	Objekt zieht sich zusammen nachdem es in Wasser gegeben wurde	Fall 2 (WF-Fall)
Typ III:	Objekt zieht sich zusammen nachdem es Temperaturen über 75°C ausgesetzt wurde	Fall 4 (FF-Fall)

Abbildung 82: Serie 3 von Beobachtungen mit drei Beobachtungstypen und die Angabe welchem Fall jeder Beobachtungstyp hinsichtlich der Disposition wasserlöslich zu sein entspricht

Die Beobachtungsserie 3 enthält die Fälle 1, 2 und 4. Das bringt uns in Schwierigkeiten. Die Beobachtungen vom Typ I (Fall 1) sprechen wieder für die Wasserlöslichkeit der Objekte des Typs A, aber aufgrund der Beobachtungen vom Typ II müssten wir Objekten des Typs A die Disposition wasserlöslich zu sein absprechen, denn die Beobachtungen vom Typ II entsprechen dem Fall 2 (siehe Abbildung 79), die Auslösebedingung ist erfüllt, die

Manifestation tritt jedoch nicht ein. Objekte, die ein solches Verhalten zeigen, die unter den Auslösebedingungen nicht die Manifestation zeigen, haben die Disposition nicht.

Diese Schwierigkeit könnten wir lösen, indem wir sagen, dass Objekte des Typs A die Disposition haben sich entweder aufzulösen *oder* sich zusammenzuziehen, wenn sie in Wasser gegeben werden. D.h., welche Disposition bestimmte Objekte haben, wird durch deren Verhalten unter bestimmten Bedingungen festgelegt. Entdecken wir beispielsweise im Anschluss an die bereits gemachten Beobachtungen (Serie 2), dass sich ein Objekt vom Typ A, nachdem es in Wasser gegeben wurde, ausdehnt hat, dann korrigieren wir unser Urteil einfach: Objekte des Typs A hätten dann die Disposition sich entweder auflösen *oder* sich zusammenzuziehen *oder* sich auszudehnen, wenn sie in Wasser gegeben werden. Diese Strategie kann natürlich leicht zu Dispositionszuschreibungen führen, die recht uninformativ und daher wenig nützlich sind. Wie soll ich mich auf das Verhalten bestimmter Objekte einstellen, wenn sie unter bestimmten Bedingungen alles Mögliche machen?

Unsere Reaktion auf die Beobachtungen (Serie 3) würde wahrscheinlich ohnehin anders ausfallen. Wir würden nämlich vermuten, dass entweder

- 1) die Auslösebedingung (ins Wasser geben) bei den Beobachtungen von Typ I und Typ II nicht gleich waren (z.B.: bei Typ I hatte das Wasser einen pH-Wert von unter 4, bei Typ II einen von über 6) oder dass
- 2) die Objekte vom Typ A nicht gleich waren, d.h. Objekte vom Typ A können in Untergruppen unterschieden werden (z.B.: bei Typ I waren es Objekte vom Typ A* und bei Typ II waren es Objekte vom Typ A**).

Trifft die Vermutung a) zu, dann sagen wir, dass Objekte vom Typ A die Disposition haben, sich in Wasser mit einem pH-Wert unter 4 aufzulösen. Trifft hingegen die Vermutung b) zu, dann sagen wir, dass Objekte vom Typ A* wasserlöslich sind.

In jedem Fall führen unsere Überlegungen dazu, dass wir Objekten eines bestimmten Typs nicht eine Disposition X zuschreiben, wenn es Objekte dieses Typs gibt, die unter den für die Disposition X relevanten Bedingungen nicht die für die Disposition X charakteristische Manifestation zeigen. Mit anderen Worten: Beobachtungen, die dem Fall 2 (siehe Abbildung 79) entsprechen sind nicht möglich, wenn Objekte eines bestimmten Typs tatsächlich eine bestimmte Disposition haben. Der Eindruck von Beobachtungen, die vermeintlich dem Fall 2 entsprechen, wird durch eine veränderte Dispositionsbeschreibung korrigiert, entweder die Disposition enthält nun eine Disjunktion oder es werden Differenzierung bei den Auslösebedingungen bzw. bei den Dispositionsträgern vorgenommen.

Wenn wir dieses Ergebnis nun auf den dispositionalen Funktionsbegriff übertragen, dann sehen wir, wieso ein dispositionaler Funktionsbegriff ein Verfehlen der Funktion durch einen Funktionsträger unmöglich macht.

Vertreter eines dispositionalen Funktionsbegriffs verstehen Funktionen als Dispositionen. So schreibt Cummins:

„Thus, function-ascribing statements imply disposition statements; to attribute a function to something is, in part, to attribute a disposition to it. If the function of x in s is to ϕ , then x has a disposition to ϕ in s. For instance, if the function of the contractile vacuole in fresh-water protozoans is to eliminate excess water from the organism, then there must be circumstances under which the contractile vacuole would actually manifest a disposition to eliminate excess water from the protozoan that incorporates it.“ (Cummins 1975, 758)

Die kontraktierende Vakuole hat Cummins zufolge eine Disposition, die zugleich ihre Funktion ist. Die Disposition besteht darin, sich ab einem bestimmten Grad an Gefülltheit (SC) mit der Zellmembran zu verbinden und so das „Abwasser“ ins Außenmedium abzugeben

(M). Wie zuvor können wir wieder die vier Kombinationen von Auslösebedingung (SC) und Manifestation (M) unterscheiden.

	Auslösebedingung (SC): Vakuole gefüllt	Manifestation (M): Vakuole entleert sich
Fall 1:	W	W
Fall 2:	W	F
Fall 3:	F	W
Fall 4:	F	F

W : erfüllt / eingetreten
F : nicht erfüllt / nicht eingetreten

Abbildung 83: Die vier möglichen Kombinationen der Ausprägungen von Auslösebedingung und Manifestation (Beispiel: Funktion der kontraktierenden Vakuole)

Auch diesmal sind die beiden letzten Fälle für die Dispositionszuschreibung und damit auch für die Funktionszuschreibung irrelevant. D.h. egal ob die Vakuole die Funktion hat sich ab einem bestimmten Füllungsgrad zu entleeren oder nicht, wenn sie nicht bis zu diesem Grad gefüllt ist, erwarten wir nicht, dass sie sich entleert. Insbesondere ist in diesen beiden Fällen auch kein Verfehlen der Funktion durch die Vakuole möglich, denn die Funktion der Vakuole ist es sich zu entleeren (M), wenn sie bis zu einem bestimmten Grad gefüllt ist (SC). Ist sie nicht bis zu einem bestimmten Grad gefüllt (SC nicht erfüllt), dann kann sie ihre Funktion *weder erfüllen noch verfehlen*.

Interessant ist hier wieder der dritte Fall, wo wir eine Entleerung beobachten, obwohl die Vakuole nicht einem bestimmten Grad an Gefülltheit erreicht hat. Intuitiv würden wir wahrscheinlich sagen, dass die Vakuole hier ihre Funktion verfehlt. Sollte sich aber hinter der Beobachtung ein ähnliches Muster verbergen, wie bei der Beobachtungsserie 2 (siehe S. 202), dann könnte man ebenso gut sagen, dass die Vakuole die Disposition und damit die Funktion hat, sich unter der Bedingung Z zu entleeren, wobei sich die Bedingung X, die in der Beobachtung vorliegt, und die Bedingung, bis zu einem bestimmten Grad gefüllt zu sein (SC), unter die Bedingung Z subsummieren lassen. In diesem Fall hätten wir aber kein Verfehlen der Funktion bei der Vakuole beobachtet, sondern ein Erfüllen der Funktion.

Damit bei der Beobachtung (Fall 3) ein Verfehlen der Funktion seitens der Vakuole vorliegt, müsste die Funktion durch ein Bikonditional ausgedrückt werden: Nur dann entleeren, wenn Vakuole gefüllt. Jetzt schreiben wir der Vakuole aber zwei Dispositionen zu²³².

Disposition 1: Wenn gefüllt (SC₁), dann entleeren (M₁)

Disposition 2: Wenn nicht gefüllt (SC₂), dann nicht entleeren (M₂)

Unter dieser Voraussetzung ist aber die Beobachtung einer Entleerung, obwohl die Vakuole einem bestimmten Grad an Gefülltheit nicht erreicht hatte, sowohl ein Fall 3 (FW-Fall) als auch ein Fall 2 (WF-Fall) (siehe Abbildung 84).

²³² Ebenso könnte man sagen, dass die Vakuole entsprechend den zwei Dispositionen auch zwei Funktionen hat. Die gemachte Beobachtung ist dann für Aussagen über die Funktion 1 irrelevant, während hinsichtlich der Funktion 2 ein Verfehlen der Funktion vorläge.

	Disposition 1		Disposition 2		
	A.-Bedingung (SC ₁): Vakuole (V) gefüllt	Manifestion (M ₁): V entleert sich	A.-Bedingung (SC ₂): V nicht gefüllt	Manifestion (M ₂): V entleert sich nicht	
Fall 1	W	W	F	F	Fall 4
Fall 2	W	F	F	W	Fall 3
Fall 3	F	W	W	F	Fall 2
Fall 4	F	F	W	W	Fall 1

W : erfüllt / eingetreten
F : nicht erfüllt / nicht eingetreten

Abbildung 84: Der Fall 3 (FW-Fall) bei Disposition 1 entspricht dem Fall 2 (WF-Fall) bei der Disposition 2 (Beispiel: Funktion der kontraktierenden Vakuole)

Ein Verfehlen der Funktion läge demnach auch bei dieser Beobachtung nur dadurch vor, dass sich unter relevanten Bedingungen (SC₂) nicht die charakteristische Manifestation zeigt (M₂), was dem Fall 2 entspricht.

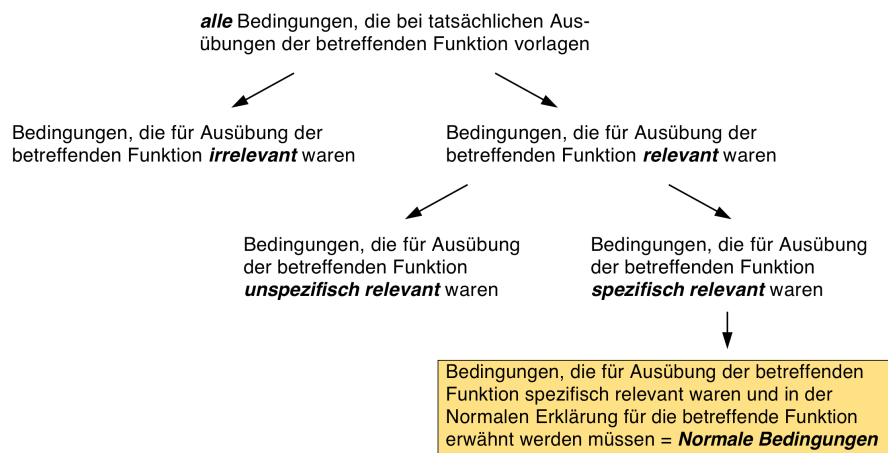
Es sind also wieder die ersten beiden Fälle, die für die Dispositions- und damit für die Funktionszuschreibung relevant sind. Wenn sich kontraktierende Vakuolen immer bei einem bestimmten Grad an Gefülltheit (SC) mit der Zellmembran verbinden und so das „Abwasser“ ins Außenmedium abgeben (M), dann haben sie die entsprechende Disposition und auch Funktion. Beobachten wir jedoch, dass eine kontraktierende Vakuole unter der genannten Bedingung (SC) die entsprechende Manifestation (M) nicht zeigt, indem sie beispielsweise platzt, nachdem sie den bestimmten Grad an Gefülltheit erreicht hatte, dann reagieren wir mit der bereits bekannten Strategie. Wir korrigieren die Dispositions- bzw. Funktionsbeschreibung, indem wir entweder eine Disjunktion in die Funktionsbeschreibung einbauen oder eine Differenzierung bei den Auslösebedingungen bzw. bei den Funktionsträgern vornehmen. Wir können also entweder behaupten, dass die Funktion der kontraktierenden Vakuolen darin besteht, sich bei einem bestimmten Grad von Gefülltheit durch Membranverschmelzung zu entleeren *oder* zu platzen. Oder wir behaupten, die Auslösebedingungen (SC) für die Entleerung mittels Membranverschmelzung sind ein bestimmter Gefülltheitsgrad der Vakuole *und* beispielsweise bestimmte Druckverhältnisse im Außenmedium, wobei Letztere bei der betreffenden Beobachtung nicht vorlagen. Oder wir behaupten, dass nur eine Untergruppe der kontraktierenden Vakuolen die besagte Funktion hat, wobei die beobachtete Vakuole aber nicht zu dieser Untergruppe gehört.

In jedem Fall ist das Resultat dieser Korrektur, dass die betreffende Beobachtung nun nicht mehr einem Fall 2, sondern einem Fall 1 oder einem Fall 4 entspricht, bzw. einer Beobachtung des falschen Objekttyps. D.h. jeder eventuell auftretende Fall 2, der dem Verfehlen der Funktion entspricht, wird in einem Fall 1 oder einem Fall 4 abgewandelt, die dem Erfüllen der Funktion bzw. einem Ereignis entsprechen, das weder ein Erfüllen noch Verfehlen der Funktion ist, oder wird in eine völlig irrelevante Beobachtung abgewandelt. Dies wiederum bedeutet, dass ein Verfehlen einer als Disposition verstandenen Funktion gar nicht möglich ist, wenn das Verfehlen darin besteht unter den relevanten Bedingungen (SC) *nicht* den für die Funktion charakteristischen Effekt (M) zu zeigen. Dies zu zeigen, war das Ziel dieses Abschnittes.

2 Millikans Auffassung des Verfehlens einer Funktion aufgrund von abNormalen Umweltbedingungen

In diesem Teilabschnitt geht es darum zu wiederholen, wie Millikan ein Verfehlen der Funktion durch einen Funktionsträger auf abNormale Umweltbedingungen zurückführt. Für

Millikans Auffassung ist dabei der Begriff der abNormalen Umweltbedingung von zentraler Bedeutung. Dieser Begriff wiederum ist in Abgrenzung zur Normalen (Umwelt)-Bedingung definiert. Normale Bedingungen, so hatten wir gesehen, sind diejenigen Bedingungen, die in der Normalen Erklärung einer Eigenfunktion *erwähnt werden müssen*. Wir hatten die Normalen Bedingungen, dann durch folgendes Schema genauer eingegrenzt (Vgl. Abschnitt 2.1.1.3.4, 73):



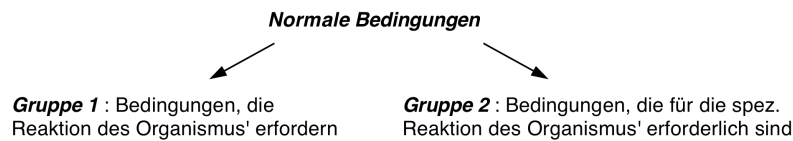
Das Verfehlen der Funktion aufgrund von abNormalen Bedingungen hatten wir uns dann anhand von drei Beispielen klargemacht, der Tarnung von Motten, der Körpertemperaturregelung durch Schweißdrüsen und der Magnetotaxis von Bakterien (Vgl. Abschnitt 2.2.2.1).

Wir hatten dabei festgestellt, dass es in allen drei Beispielen Bedingungen (Fressfeind anwesend, überhöhte Körpertemperatur, Sauerstoff in der Umgebung) gibt, auf die der jeweilige Organismus (Motte, Mensch, Bakterium) reagieren muss, weil er sonst nicht überlebt. Es drohte: Gefressenwerden, Hitzetod, tödliche Vergiftung. Alle Organismen haben nun einen Funktionsträger (Außenhülle, Schweißdrüse, Magnetosom) dessen Aufgabe es ist, das jeweils drohende Resultat zu verhindern. Der Funktionsträger erfüllt seine Funktion also, wenn er einen bestimmten Effekt (Nichtgefressenwerden, Absenken der Körpertemperatur, Vermeidung von sauerstoffreichem Wasser) bewirkt. Diesen Effekt wiederum bewirkt der Funktionsträger auf eine bestimmte Weise (durch Tarnung, durch Schweißabsonderung, durch Schwimmen zum geomagnetischen Norden). Damit der Effekt wiederum auf die bestimmte Weise erreicht werden kann, müssen weitere Bedingungen (Motte sitzt auf dunklem Untergrund, die Luftfeuchtigkeit liegt unter 100%, das Bakterium befindet sich in der nördlichen Hemisphäre) erfüllt sein.

Dies hatte uns dann dazu bewogen zwei Gruppen von Normalen Bedingungen zu unterscheiden:

Zur *ersten Gruppe* gehören diejenigen Bedingungen, aufgrund derer überhaupt ein Handlungsbedarf seitens des Organismus besteht.

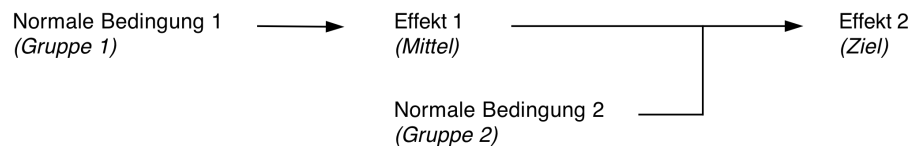
Zur *zweiten Gruppe* gehören diejenigen Bedingungen, die für die spezifische Problemlösung des Organismus' gegeben sein müssen.



Unsere Ergebnisse hatten wir dann in der folgender Tabelle und folgendem Schema zusammengefasst (Vgl. Abschnitt 2.2.2.1, S. 113f).

Beispiel	Normale Bedingungen		Effekte	
	Gruppe 1	Gruppe 2	(Mittel)	(Ziel)
Motte (dunkel)	Freßfeind anwesend	Motte auf dunklem Untergrund	kein Kontrast zur Umgebung (Tarnung)	nicht gefressen werden
Schweißdrüse	erhöhte Körpertemperatur	Luftfeuchtigkeit unter 100%	Schweiß absondern	Körpertemperatur absenken
Bakterium (Magnetotaxis)	O ₂ -reiches Oberflächenwasser	Bakterium auf nördl. Hemisphäre	Schwimmen zum magnetischen Norden	Vermeiden von O ₂ -reichem Wasser

Schema, das den Wirkungszusammenhang zwischen den zwei Gruppen von Normalen Bedingungen und den verschiedenen Effekten (Mittel/Ziel) zeigt:



Insgesamt hatte wir festgestellt, dass die Beispiele nahelegen, dass nur ein Abweichen von der zweiten Gruppe von abNormalen Bedingungen überhaupt den Eindruck vermitteln kann, dass die jeweiligen Funktionsträger in diesen Situationen ihre Funktion verfehlen. Deshalb wollten wir auch Millikan so verstehen, dass sie sich bei einem Verfehlen aufgrund von abNormalen Umweltbedingungen immer auf ein Abweichen von der zweiten Gruppe von Normalen Bedingungen bezieht.

Millikans Erklärung für ein Verfehlen der Funktion aufgrund von abNormalen Umweltbedingungen hatten wir dann so zusammengefasst:

Das Verfehlen der Funktion aufgrund von abNormalen Umweltbedingungen entsteht dadurch, dass der Funktionsträger, dessen Aufgabe es ist ein Resultat, das die Folge einer bestimmten Bedrohung ist, zu verhindern, nicht in der Lage ist das Resultat bei Vorliegen der Bedrohung tatsächlich zu verhindern. Es misslingt ihm, nicht weil er defekt ist, sondern weil andere Bedingungen, die für seine spezifische Arbeitsweise notwendig sind und sich seiner Kontrolle entziehen, nicht erfüllt sind.

Dementsprechend sollte für Millikan gelten:

- (1) Die dunkle Außenhülle einer Motte verfehlt ihre Funktion, wenn die Motte auf hellem Untergrund sitzend von einem Fressfeind entdeckt und gefressen wird.
- (2) Eine Schweißdrüse verfehlt ihre Funktion, wenn sie bei 100% Luftfeuchtigkeit keine Absenkung der erhöhten Körpertemperatur durch Schweißabsonderung bewirkt.
- (3) Die Magnetosome eines Bakteriums verfehlen ihre Funktion, wenn sie auf der südlichen Hemisphäre das Bakterium nicht vom sauerstoffreichen Oberflächenwasser fernhalten, indem sie es Richtung geomagnetischer Norden schwimmen lassen.

Wir haben nun hinreichend genau wiederholt, wie Millikan das Verfehlen der Funktion durch einen Funktionsträger auf abNormale Umweltbedingungen zurückführt, und damit das Ziel dieses Teilabschnittes erreicht.

3 Was die Erklärung des Verfehlens der Funktion aufgrund von abNormalen Umweltbedingungen anbelangt, ist Millikans Theorie in derselben Situation wie der von ihr kritisierte dispositionale Funktionsbegriff

In diesem Teilabschnitt versuche ich zu zeigen, dass Millikans Erklärung des Verfehlens einer Funktion aufgrund von abNormalen Umweltbedingungen in derselben Situation wie der von ihr kritisierte dispositionale Funktionsbegriff ist. Gelingt dies, dann folgt daraus²³³, dass sich auch im Rahmen von Millikans Biosemantik das Verfehlen einer Funktion aufgrund von abNormalen Umweltbedingungen und damit Missrepräsentationen bei Organismen mit intakten kognitiven Systemen nicht erklären lassen.

Auch bei Millikan sind Funktionen relativ zu bestimmten Bedingungen bestimmt, eben den Normalen Bedingungen. D.h. ein Funktionsträger erfüllt seine Funktion, wenn er wie seine Vorfahren unter den gleichen Bedingungen den gleichen Effekt zeitigt, der zur weiteren Reproduktion der Funktionsträger geführt hat. Indem wir uns nun aber lediglich auf einen Teil dieser Normalen Bedingungen konzentrieren, nämlich auf diejenigen, aufgrund derer überhaupt ein „Handlungsbedarf“ seitens des Organismus besteht (erste Gruppe von Normalen Bedingungen), ergeben sich Fälle, in denen diese Normalen Bedingungen erfüllt sind, aber der Effekt nicht erreicht wird. Die Körpertemperatur ist erhöht, aber die Schweißdrüse erreicht kein Absenken der Körpertemperatur (trotz Schweißabsonderung). Diesen Fall können wir in unserem bewährten Schema so darstellen:

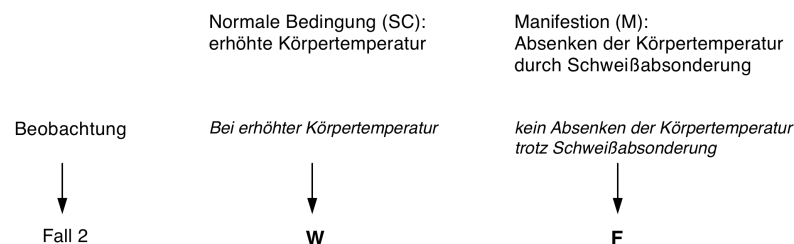


Abbildung 85: Scheinbarer WF-Fall (Beispiel: Schweißdrüse)

Dieser Fall entspricht also scheinbar genau einem jener Fälle, die es bei einem dispositionalen Funktionsbegriff gar nicht geben konnte, also einem Fall 2 oder, wie wir auch sagen können, einem WF-Fall. Dieser Eindruck täuscht jedoch, wie wir uns wiederum an Millikans Schweißdrüsen-Beispiel klar machen können. Tatsächlich müssen wir nämlich zwei Bedingungen²³⁴ und zwei Effekte unterscheiden.

Die beiden Bedingungen sind:

R = erhöhte Körpertemperatur
 S = Luftfeuchtigkeit unter 100%

Die beiden Effekte sind:

X = Schweiß absondern
 Y = Körpertemperatur absenken

²³³ Unter der bereits erwähnten und von Millikan geteilten Voraussetzung, dass dispositionale Funktionsbegriffe kein Verfehlen der Funktion zulassen.

²³⁴ Die zugleich die Normalen Bedingungen sind.

Die Funktion der Schweißdrüse ist bei Millikan so bestimmt, dass diese *beide* Effekte (*X und Y*) bewirkt²³⁵, wenn *beide* Bedingungen (*R und S*) gegeben sind. Dies ist auch zentral für die (most proximate) Normale Erklärung der Eigenfunktion der Schweißdrüse. Wir können jetzt die Funktion der Schweißdrüse als die Wenn-Dann-Beziehung, wenn *T*, dann *Z* wiedergeben, indem wir die Konjunktion von *R und S* zu *T* zusammenfassen und die Konjunktion von *X und Y* zu *Z*.

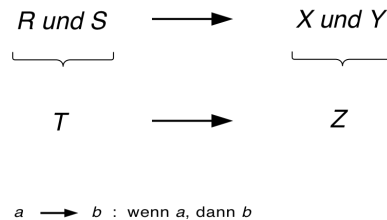


Abbildung 86: Die Eigenfunktion der Schweißdrüse als Wenn-Dann-Beziehung: *Wenn* Körpertemperatur erhöht (R) *und* Luftfeuchtigkeit unter 100% (S), *dann* Schweiß absondern (X) *und* Körpertemperatur senken (Y)

Wenn wir nun wie zuvor bei unseren Überlegungen zur Dispositionszuschreibung alle möglichen Kombinationen der Ausprägungen von Normalen Bedingungen (erfüllt/nicht erfüllt) und Effekten (eingetreten/nicht eingetreten) auflisten, dann erhalten wir genau 16 Fälle.

	Auslösebedingung (SC):		Manifestion (M):	
	Körpertemperatur erhöht	Luftfeuchtigkeit unter 100%	Schweiß absondern	Körpertemperatur absenken
Fall 01:	W	W	W	W
Fall 02:	W	F	W	W
Fall 03:	F	W	W	W
Fall 04:	F	F	W	W
Fall 05:	W	W	F	W
Fall 06:	W	F	F	W
Fall 07:	F	W	F	W
Fall 08:	F	F	F	W
Fall 09:	W	W	W	F
Fall 10:	W	F	W	F
Fall 11:	F	W	W	F
Fall 12:	F	F	W	F
Fall 13:	W	W	F	F
Fall 14:	W	F	F	F
Fall 15:	F	W	F	F
Fall 16:	F	F	F	F

W : erfüllt / eingetreten
F : nicht erfüllt / nicht eingetreten

Abbildung 87: Die 16 möglichen Kombinationen der Ausprägungen von Auslösebedingungen und Manifestationen (Beispiel: Funktion der Schweißdrüse)

²³⁵ Genaugenommen: Y durch X.

Der 10. Fall ist derjenige, in dem nach Millikan die Schweißdrüse ihre Funktion aufgrund abNormaler Umweltbedingungen verfehlt. Wenn wir nun aber die Eigenfunktion der Schweißdrüse so verstehen, wie wir sie eben beschrieben haben, nämlich als ein Bewirken des Effektes *Z* (*X und Y*) unter der Bedingung *T* (*R und S*), dann ist dieser Fall kein Beispiel für ein Ausbleiben des Effektes unter den relevanten Bedingungen, kein WF-Fall, sondern ein Beispiel für das Ausbleiben des Effektes, wenn die relevanten Bedingungen nicht erfüllt sind, also ein FF-Fall. Mit anderen Worten: Eine Schweißdrüse verfehlt ihre Funktion, das Absenken der Körpertemperatur durch Schweißabsonderung (Effekt *Z*), nicht, wenn der Effekt *Z* ausbleibt, weil die dafür relevante Bedingung (*T*) nicht erfüllt ist²³⁶. Ein Ding, das (nur) unter bestimmten Bedingungen etwas bewirken soll, versagt nicht, wenn diese Bedingungen nicht vorliegen und es dann seine spezifischen Wirkungen nicht erzielt. Generell gilt, wenn die Normalen Bedingungen in der oben genannten Weise in die Bestimmung der Eigenfunktion mit eingehen, d.h. wenn die Eigenfunktion eines Funktionsträgers darin besteht, unter Normalen Bedingungen einen bestimmten Effekt zu erzeugen, dann kann es kein Verfehlen der Funktion unter abNormalen Bedingungen geben. Dies gilt zumindest solange, wie wir ein Verfehlen der Funktion so verstehen, wie wir es bei der Kritik des dispositionalen Funktionsbegriffs herausgearbeitet haben, nämlich als ein Ausbleiben des relevanten Effekts *unter den relevanten Bedingungen*. Wir könnten dieses Verständnis aber auch aufgeben und ein Verfehlen einer Funktion so verstehen, wie dies Millikan im Fall der Schweißdrüse andeutet hat. Bei dieser Wahl bricht aber auch die Kritik am dispositionalen Funktionsbegriff zusammen, wie sich leicht zeigen lässt. Angenommen ich schreibe einem Funktionsträger die als Disposition verstandene Funktion zu unter den Bedingungen *A* und *B* den Effekt *E* zu zeigen. Von einem solchen Funktionsträger könnte ich dann zu Recht behaupten, dass er seine Funktion verfehlt, wenn er unter den Bedingungen *A* und *nicht B* (!) den Effekt *E* nicht bewirkt. In diesem Sinne ist Millikans Auffassung vom Verfehlen einer Funktion aufgrund von abNormalen Umweltbedingungen mit einem dispositionalen Funktionsbegriff kompatibel. Dies galt es in diesem Abschnitt zu zeigen.

Insgesamt folgt daraus, dass sich im Rahmen von Millikans Biosemantik das Verfehlen einer Funktion aufgrund von abNormalen Umweltbedingungen und damit Missrepräsentationen bei Organismen mit intakten kognitiven Systemen nicht erklären lassen, zumindest wenn man wie Millikan voraussetzt, dass dispositionale Funktionsbegriffe kein Verfehlen der Funktion zulassen.

²³⁶ Die Bedingung *T* ist in diesem Fall nicht erfüllt, weil ein Teil der Konjunktion, nämlich Bedingung *S*, nicht erfüllt ist.

Fazit

Die zentrale **Frage** dieser Arbeit war:

Kann Millikans Biosemantik das Problem der Missrepräsentation lösen?

Meine abschliessende **Antwort** ist insgesamt positiv: Millikans Biosemantik enthält tatsächlich eine Lösung für das Problem der Missrepräsentation. Diese Lösung ist allerdings in bedeutsamer Weise eingeschränkt.

Zunächst hatten wir im zweiten Teil der Arbeit ausführlich dargelegt, dass Millikan genaugenommen zwei Lösungsansätze anbietet.

1. Missrepräsentationen aufgrund von abNormalen Umweltbedingungen (Vgl. Abschnitt 2.2.2)
2. Missrepräsentationen aufgrund von missgebildeten Funktionsträgern (Vgl. Abschnitt 2.2.3)

Es ist ein wichtiger Beitrag dieser Arbeit, diese beiden Ansätze klar auseinanderzuhalten. Damit trägt die Arbeit grundlegend zu mehr Transparenz bei der Bewertung und Analyse der biosemantischen Lösung(en) des Problems der Missrepräsentation bei.

Im dritten Teil der Arbeit sind wir dann kritischen Einwänden gegen Millikans Lösungen nachgegangen. Einige der Einwände konnten bei genauerer Betrachtung entkräftet werden. Dies betraf die Einwände von Fodor (1990), Pietroski (1992), Keil (1993), Perlman (2000), Mendola (2008) und Stegmann (2009). Andere Einwände (Davies 1994, Nanay 2010, 2012) konnten nicht oder nur teilweise entkräftet werden, woraus sich erhebliche Einschränkungen für Millikans Lösungsansätze ergeben.

Aus Davies (1994) Einwand ergab sich, dass Millikans Lösungen vorerst keine Geltung im Bereich von Organismen haben, die sich sexuell fortpflanzen. Eine Ausdehnung der Geltung in diesen Bereich scheint möglich, ist aber bisher nicht geleistet worden.

Aus Nanays (2010, 2012) Einwand ergab sich, dass Millikans Lösungsansatz, Missrepräsentationen aufgrund von missgebildeten Funktionsträgern zu erklären, nur für nachträglich deformierte Funktionsträger konsistent ist. Missbildungen im eigentlichen Sinne können in Millikans Theorie keine Missrepräsentation hervorrufen.

Aus meinem eigenen Einwand ergab sich, dass Millikans Lösungsansatz, Missrepräsentationen aufgrund von abNormalen Umweltbedingungen zu erklären, insgesamt nicht konsistent ist. Bei völlig intakten Organismen kann es also im Rahmen von Millikans Theorie keine Missrepräsentationen geben.

Diese Einschränkungen sollten uns bezüglich der Rolle, die Millikans Theorie für die Kognitionswissenschaften als Antwort auf den Homunkuluseinwand spielen könnte, skeptisch machen. Mit anderen Worten: Die Absicherung der Auffassung von Kognition als Informationsverarbeitung durch Millikans Theorie ist aufgrund der genannten Einschränkungen problematisch, weil ein grosser Teil dessen, worüber die Kognitionswissenschaften sprechen wollen, von Millikans Theorie nicht abgesichert werden kann.

Auf der anderen Seite darf meines Erachtens auch die positive Seite des Resultats, nämlich dass es überhaupt eine Lösung des Problems der Missrepräsentation gibt, nicht unterschätzt werden. Das Lösen dieses Problems gilt vielen Philosophen nach wie vor als Quadratur des Kreises. Millikans Theorie stellt damit einen wichtigen Fortschritt dar und ihre Theorie kann als Ausgangspunkt für naturalistische Theorien des Gehalts dienen, die tatsächlich den Homunkuluseinwand gegenüber den Kognitionswissenschaften entkräften können.

Literatur

Adams, Fred and Ken Aizawa 2010: „Causal Theories of Mental Content“, The Stanford Encyclopedia of Philosophy (Spring 2010 Edition), Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/spr2010/entries/content-causal/>>.

Adriaans, Pieter 2013: „Information“, The Stanford Encyclopedia of Philosophy (Fall 2013 Edition), Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/fall2013/entries/information/>>.

Agar, N. 1993: „What do Frogs Really Believe?“ *Australasian Journal of Philosophy* 71, 1–12.

Anscombe, G. E. M. 1963 [1957]: *Intention* (Cambridge, Mass.: Harvard University Press), 2. Aufl.

Baker, Lynne R. 1995: *Explaining Attitudes. A Practical Approach to the Mind* (Cambridge, Mass.: Cambridge University Press).

Baker, M., P. Wolanin and J. Stock 2005: „Signal transduction in bacterial chemotaxis“ *Bioessays* 28, 9–22.

Beckermann, A. 2008: *Analytische Einführung in die Philosophie des Geistes* (Berlin: de Gruyter), 3. Aufl.

Bechtel, W. 2008: *Mental mechanisms: Philosophical perspectives on cognitive neuroscience* (London: Routledge).

Bechtel, William und Robert C. Richardson 1993: *Discovering complexity. decomposition and localization as strategies in scientific research* (Princeton, New Jersey : Princeton University Press).

Bechtel, W., Adele Abrahamsen und George Graham 1999: „The Life of Cognitive Science“ in: Bechtel, W. und George Graham (eds.) *A Companion to Cognitive Science* (Malden, Mass: Blackwell), 1–104.

Bechtel, W. und George Graham (eds.) 1999: *A Companion to Cognitive Science* (Malden, Mass: Blackwell).

Bennett, Maxwell and Peter Hacker 2007: „The Conceptual Presuppositions of Cognitive Neuroscience: A Reply to Critics“ in: Bennett, Maxwell et al. *Neuroscience and philosophy : brain, mind, and language* (New York: Columbia University Press), 127–162.

Bermúdez, José Luis 2005: *Philosophy of Psychology. A Contemporary Introduction* (New York: Routledge).

Bernard et al. 2012: „Ultraviolet radiation damages self noncoding RNA and is detected by TLR3“ *Nature Medicine* 18, 1286–1290.

Bickhard, Mark H. and Lore Terveen 1995: *Foundational issues in artificial intelligence and cognitive science. impasse and solution* (Amsterdam: Elsevier).

Biletzki, Anat and Anat Matar 2014: "Ludwig Wittgenstein", The Stanford Encyclopedia of Philosophy (Spring 2014 Edition), Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/spr2014/entries/wittgenstein/>>

Bischof, Norbert 1998: *Struktur und Bedeutung. Eine Einführung in die Systemtheorie für Psychologen* (Bern: Huber).

Bischof, Norbert 2009: *Psychologie. Ein Grundkurs für Anspruchsvolle* (Stuttgart: Kohlhammer).

Botterill, George and Peter Carruthers 2003: *The Philosophy of Psychology* (Cambridge University Press (Virtual Publishing)).

Carnap Rudolf 1947: *Meaning and Necessity* (Chicago: The University of Chicago Press).

Carroll, John W. 2012: „Laws of Nature“, The Stanford Encyclopedia of Philosophy (Spring 2012 Edition), Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/spr2012/entries/laws-of-nature/>>.

Carroll, Sean B. 2007: *Endless Forms Most Beautiful* (London: Phoenix).

Cassam, Quassim 2007: „The Possibility of Knowledge“ *Grazer Philosophische Studien* 74, 125-141.

Chemero, Anthony 2009: *Radical embodied cognitive science* (Cambridge, Mass.: MIT Press).

Clark, Andy 2008: *Supersizing the mind : embodiment, action, and cognitive extension* (Oxford: Oxford University Press).

Choi, S. und M. Fara 2012: „Dispositions“, The Stanford Encyclopedia of Philosophy (Spring 2012 Edition) Zalta, E. (Hrsg.). URL = <<http://plato.stanford.edu/archives/spr2012/entries/dispositions/>>.

Christensen, Wayne D. and Cliff A. Hooker 2004: „Representation and the Meaning of Life“ in: Clapin, Hugh, Phillip Staines and Peter Slezak (eds.) *Representation in Mind. New Approaches to Mental Representation* (Amsterdam: Elsevier), 41-70.

Clapin, H. et al. (ed.) 2004: *Representation in Mind: New Approaches to Mental Representation* (Amsterdam: Elsevier).

Crane, Tim 2001: *Elements of Mind. An Introduction to the Philosophy of Mind* (Oxford: Oxford University Press).

Craver, Carl. 2009 [2007]: *Explaining the Brain. Mechanisms and the Mosaic Unity of Neuroscience* (Oxford: Oxford University Press).

Crick, Francis H. C. 1955: „On Degenerate Templates and the Adaptor Hypothesis“, unveröffentlichte Note an den RNA Tie Club, nicht datiert, Mitte Jan. 1955.

Cummins, R. 1975: „Functional Analysis“ *Journal of Philosophy* 72, 741–765.

Cummins, Robert and Martin Roth 2010: „Meaning and Content in Cognitive Science“ in: Robert Cummins *The World in the Head* (Oxford: Oxford University Press), 174-193.

Davidson, Donald (1990) [1975]: „Denken und Reden“ in: ders. *Wahrheit und Interpretation* (Frankfurt am Main: Suhrkamp), 224-246.

Davies, Paul Sheldon 1994: „Troubles For Direct Proper Functions“ *Noûs* Vol. 28, No. 3, 363-381.

Dawkins, Richard 2000: *Der erweiterte Phänotyp. Der lange Arm der Gene* (Heidelberg : Spektrum Akademischer Verlag) [engl. Original: Dawkins, Richard 1982: *The extended phenotype : the gene as the unit of selection* (Oxford/San Francisco: Freeman)].

Deacon, Terrence W. 2010: „What is missing from theories of information?“ in: Paul Davies and Niels Henrik Gregersen (eds.) *Information and the Nature of Reality. From Physics to Metaphysics* (Cambridge: Cambridge University Press), 146-169.

Dennett, Daniel 1996: *Kinds of Minds. Toward an Understanding of Consciousness* (New York: BasicBooks).

Dennett, Daniel 1969: *Content and Consciousness* (London: Routledge).

Dennett, Daniel 1978: *Brainstorms. Philosophical Essays on Mind and Psychology* (Montgomery - Vt.: Bradford Books).

Dennett, Daniel 1987: *The Intentional Stance* (Cambridge: MIT Press).

Dennett, Daniel 1995: *Darwin's Dangerous Idea. Evolution and the Meanings of Life* (London: Penguin).

Detel, W. 2001a: „Haben Frösche und Sumpfmenschen Gedanken? Einige Probleme der Teleosemantik.“ *Deutsche Zeitschrift für Philosophie* 49, 601-626.

Detel, W. 2001b: „Teleosemantik. Ein neuer Blick auf den Geist?“ *Deutsche Zeitschrift für Philosophie* 49, 465-491.

Dray, William H. 1966 [1957]: *Laws and Explanation in History* (Oxford: Clarendon Press).

Dretske, Fred 1981: *Knowledge & the flow of information* (Cambridge, Mass.: MIT Press).

Dretske, Fred 1986: „Misrepresentation“ in: Radu, B. (ed.) *Belief: Form, content and function* (Oxford: Clarendon Press), 17-36.

Dretske, Fred 1988: *Explaining Behavior* (Cambridge, Mass.: MIT Press).

Dretske, Fred 2009. „Information-theoretic semantics“ in: McLaughlin, B., A. Beckermann und S. Walter (eds.) *The Oxford Handbook of Philosophy of Mind* (Oxford: Clarendon Press), 381-393.

Eells, Ellery 1991: *Probabilistic Causality* (Cambridge, U.K.: Cambridge University Press).

Eigen, Manfred und Ruthild Winkler 1978 [1975]: *Das Spiel* (München: Piper).

- Eigen, Manfred 1992: *Stufen zum Leben* (München: Piper), 2. Aufl.
- Evans, Gareth 1982: *The varieties of reference* (Oxford : Clarendon Press).
- Fodor, J. A. 1984: „Semantics, Wisconsin Style“ *Synthese* 59, 231–250.
- Fodor, J. A. 1990: *A theory of content and other essays* (Cambridge, Mass.: MIT Press).
- Frege, Gottlob 1990 [1892]: „Über Sinn und Bedeutung“ in: Ignacio Angelelli (ed.) *Gottlob Frege. Kleine Schriften* (Hildesheim: Olms), 143-162.
- Frege, Gottlob 1976 [1891]: „Brief an Husserl 24.5. 1891“ in: Gottfried Gabriel (ed.) *Gottlob Frege. Wissenschaftlicher Briefwechsel* (Hamburg: Meiner), 94-98.
- Frith, Chris 2007: *Making up the Mind. How the Brain Creates our Mental World* (Singapore: Blackwell).
- Gallagher, Shaun and Dan Zahavi 2008: *The Phenomenological Mind. An Introduction to Philosophy of Mind and Cognitive Science* (New York: Routledge).
- Gallistel, R. 2003: „Learning and Representation“ in: R. Menzel (ed.) *Learning Theory and Behavior. Vol. [1] of Learning and Memory: A Comprehensive Reference*, 4 vols. (J.Byrne Editor) (Oxford: Elsevier), 227-242.
- Gallistel, C.R. 2006: „The nature of learning and the functional architecture of the brain“ in: Q. Jing, et al. (eds) *Psychological Science Around the World, vol 1. Proceedings of the 28th International Congress of Psychology* (Sussex: Psychology Press), 63-71.
- Gamow, George 1954: “Possible Relation between Deoxyribonucleic Acid and Protein Structures” *Nature* 173, 318.
- Garner, Wendell R. 1988: „The contribution of information theory to psychology“ in: William Hirst (ed.) *The making of cognitive science : essays in honor of George A. Miller* (Cambridge: Cambridge University Press), 19-35.
- Giere, R.N. 2000: „Theories“ in: W.H. Newton-Smith (ed.) *A Companion to the Philosophy of Science* (Malden, Mass.: Blackwell), 515-524 .
- Glock, Hans-Johann 2000: „Animals, Thoughts and Concepts” *Synthese* 123, 35-64.
- Glock, Hans-Johann 2013: „Animal Minds: Philosophical and Scientific Aspects“ in: T.P. Racine and K.L. Slaney (eds.) *A Wittgensteinian Perspective on the Use of Conceptual Analysis in Psychology* (Basingstoke: Palgrave Macmillan), 130-152.
- Glock, Hans-Johann 2014: „Propositional Attitudes, Intentional Contents and other Representationalist Myths“ in: Coliva, A.; Moya-Sharrock, D; Munz, V. (eds.) *Mind, Language and Action* (New York: de Gruyter), 512-537.
- Glock, Hans-Johann (forthcoming): *Animal Minds: conceptual problems*.

Goetz, Aaron T., Todd K. Shakelford, Steven M. Platek 2009: „Introduction to evolutionary psychology: A Darwinian approach to human behavior and cognition“ in: Steven M. Platek and Todd K. Shakelford (eds.) *Foundations in Evolutionary Cognitive Neuroscience* (Cambridge: Cambridge University).

Godfrey-Smith, Peter 1989: „Misinformation“ *Canadian Journal of Philosophy* 19, 533-550.

Godfrey-Smith, Peter 1996: *Complexity and the Function of Mind in Nature*. Cambridge Studies in Philosophy and Biology (Cambridge & New York: Cambridge University Press).

Godfrey-Smith, Peter 2003: *Theory and reality : an introduction to the philosophy of science* (Chicago: University of Chicago Press).

Godfrey-Smith, Peter 2007: „Information in Biology“ in: Hull, David L. & Michael Ruse (eds.) *The Cambridge Companion to the Philosophy of Biology* (Cambridge: Cambridge University Press), 103-119.

Godfrey-Smith, Peter 2013: „Signals, Icons, Beliefs“ in: Kingsbury, J. , D. Ryder und K. Williford (eds.) *Millikan and Her Critics* (Oxford: Blackwell), 41-62.

Godfrey-Smith, Peter 2014: *Philosophy of Biology* (Princeton and Oxford: Princeton University Press).

Goldblum, Naomi 2004: *The Brain-Shaped Mind. What the Brain Can Tell Us about the Mind* (Cambridge: Cambridge University Press).

Goode R., P. E. Griffiths 1995: „The Misuse of Sober's Selection for/Selection of Distinction“ *Biology and Philosophy* 10, 99-108.

Goodman, Nelson 1983 [1954]: *Fact, Fiction, and Forecast* (Cambridge, MA: Harvard University Press), 4th ed.

Gregory, R. L. 1961: „The Brain as Engineering Problem“ in: Thorpe, W. H. und O. L. Zangwill (eds.) *Current Problems in Animal Behavior* (Cambridge: Cambridge University Press), 307-330.

Grice, P. 1989: „The Causal Theory of Perception“, ders. *Studies in the Way of Words* (Harvard: Harvard University Press), 224–247.

Griffith, Paul und Karola Stotz 2013: *Genetics and Philosophy. An Introduction* (Cambridge: Cambridge University Press).

Grossenbacher, Peter G. 2001: „A phenomenological introduction to the cognitive neuroscience of consciousness“ in: Grossenbacher, Peter G. (ed.) *Finding consciousness in the brain : a neurocognitive approach* (Amsterdam: John Benjamins), 1-19.

Habermas, Jürgen 1981: *Theorie des kommunikativen Handelns* Bd. 1 (Frankfurt am Main : Suhrkamp Verlag).

Hall, Richard J. 1990: „Does Representational Content Arise from Biological Function?“ *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association* Vol. 1990, Volume One: Contributed Papers, 193-199.

Harré, Rom 2002: *Cognitive science. A philosophical introduction* (London: Sage).

Haugeland, John 1998: *Having Thought. Essays in the metaphysics of mind* (Cambridge, Mass.: Harvard University Press).

HBP Report, April 2012.

(https://www.humanbrainproject.eu/documents/10180/17648/TheHBPReport_LR.pdf/18e5747e-10af-4bec-9806-d03aead57655, 12.05.2015, 17:00)

Heil, John 2004: *Philosophy of Mind. A contemporary introduction* (New York: Routledge).

Hempel, Carl Gustav 1977: *Aspekte Wissenschaftlicher Erklärung* (Berlin: de Gruyter).

Henderson, Harry 2007: *Artificial intelligence. mirrors for the mind* (New York: Chelsea House).

Hiley, B. J. 2004: „Information, quantum theory and the brain" in: Globus, Gordon G., Karl H. Pribram and Giuseppe Vitiello (eds.) *Brain and Being. At the boundary between science, philosophy, language and arts* (Amsterdam: John Benjamins), 199-216.

Hitchcock, Christopher 1996: „A probabilistic theory of second order causation" *Erkenntnis* 44, Issue 3, 369-377.

Hofbauer, J. und K. Sigmund 1998: *Evolutionary Games and Population Dynamics* (Cambridge: Cambridge University Press).

Hull, David L. und Michael Ruse (eds.) 2007: *The Cambridge Companion to the Philosophy of Biology* (Cambridge: Cambridge University Press).

Hunt, Earl 2005: „Information Processing and Intelligence: Where We Are and Where We Are Going" in: Sternberg, Robert J. und Jean E. Pretz (eds.) *Cognition and Intelligence. Identifying the Mechanisms of the Mind* (New York: Cambridge University Press), 1-25.

Ismael, J. T. 2007: *The Situated Self* (Oxford: Oxford University Press).

Jacob, P. 1997: *What Minds Can Do: Intentionality in a Non-Intentional World* (Cambridge: Cambridge University Press).

Janich, Peter 1999: *Die Naturalisierung der Information* (Stuttgart: Franz Steiner Verlag).

Kahneman, D., P. Slovic, and A. Tversky (eds.) 1982: *Judgement under Uncertainty: Heuristics and Biases* (Cambridge: Cambridge University Press).

Kalish, Donald 1972: „Semantics" in: Edwards, Paul (ed.) *The encyclopedia of philosophy* (New York : Macmillan & Free Press), 348-358.

Kandel, Eric 1976: *Cellular Basis of Behavior* (San Francisco: W.H. Freeman and Company).

Kandel, Eric 2006: *In search of memory: the emergence of a new science of mind* (New York: Norton).

Kandel et al. (eds.) 2011: *Neurowissenschaften. Eine Einführung* (Heidelberg: Spektrum Akademischer Verlag).

Kant, Immanuel 2001 [1790]. *Kritik der Urteilkraft* (Hamburg: Meiner).

Keil, Geert 1993a: *Kritik des Naturalismus* (Berlin [u.a.]: de Gruyter).

Keil, Geert 1993b: „Biosemantik: ein degenerierendes Forschungsprogramm?“ in: Lenk, Hans und Hans Poser (eds.) *Neue Realitäten - Herausforderung der Philosophie* (Berlin: Akademie Verlag), 86-93.

Keil, Geert 2003: „Homunkulismus in den Kognitionswissenschaften“ in: Köhler, Wolfgang R. und Hans-Dieter Mutschler (eds.) *Ist der Geist berechenbar? Philosophische Reflexionen* (Darmstadt: Wissenschaftliche Buchgesellschaft), 77-112.

Keil, Geert 2000: *Naturalismus und Intentionalität* in: Keil, Geert und Herbert Schnädelbach (eds.) *Naturalismus. Philosophische Beiträge* (Frankfurt am Main : Suhrkamp), 187-204.

Kenny, Anthony 1991 [1971]: „The Homunculus Fallacy“ in: Hyman, John (ed.) *Investigating Psychology: Sciences of the Mind after Wittgenstein* (London: Routledge), 155-165.

Kettlewell, Henry Bernard Davis 1955: „Selection experiments in industrial melanism in the Lepidoptera“ *Heredity* 9, 323-342.

Kirschner, W. Marc und John C. Gerhart 2007: *Die Lösung von Darwins Dilemma* (Reinbek: Rowohlt Taschenbuch Verlag).

Kriegel, Uriah 2009: *Subjective Consciousness. A Self-Representational Theory* (Oxford: Oxford University Press).

Künne, Wolfgang 1972: „Edmund Husserl: Intentionalität“ in: Speck, J. (ed.) *Grundprobleme der großen Philosophen* (Göttingen: Vandenhoeck und Ruprecht), 165-215.

Küppers, Bernd-Olaf 1983: *Molecular Theory of Evolution: Outline of a Physico-Chemical Theory of the Origin of Life* (Berlin: Springer).

Küppers, Bernd-Olaf 1986: *Der Ursprung biologischer Information: zur Naturphilosophie der Lebensentstehung* (München: Piper).

Kurthen, Martin 1992: *Neurosemantik* (Stuttgart: Enke).

Larkin, J. H., H. A. Simon (1987): „Why a diagram is (sometimes) worth ten thousand words“ *Cognitive Science* 11, 65-99.

Laubichler, Manfred D. 2007: „Evolutionary Developmental Biology“ in: Hull, David L. und Michael Ruse (eds.) *The Cambridge Companion to the Philosophy of Biology* (Cambridge: Cambridge University Press), 342-360.

Laurence, Stephen und Eric Margolis 2003: „Concepts and Conceptual Analysis“ *Philosophy and Phenomenological Research* Vol. 67, No. 2, 253-282.

Lenk, Hans 2001: *Das Denken und sein Gehalt* (München: Oldenbourg).

- Lewis, David 1975: *Konvention: Eine sprachphilosophische Abhandlung* (Berlin: de Gruyter) [engl. Original: David Lewis 1969: *Conventions* (Cambridge, Mass.: Harvard University Press)].
- Lorenz, Kuno 1984: „Semantik“ in: Jürgen Mittelstraß (ed.) *Enzyklopädie Philosophie und Wissenschaftstheorie* (Mannheim: Bibliographisches Institut), 768-775.
- Lowe, E. L. 2000: *An Introduction to the Philosophy of Mind* (Cambridge: Cambridge University Press).
- Lyre, Holger 2002: *Informationstheorie. Eine philosophisch-naturwissenschaftliche Einführung* (München: Fink).
- Macdonald, G. 1995: „Introduction: The Biological Turn“ in: Macdonald, C. und G. Macdonald (eds.) *Philosophy of psychology* (Oxford UK; Cambridge Mass.: Blackwell), 238-252.
- Macdonald, G. und D. Papineau. 2006. „Introduction: Prospects and Problems for Teleosemantics“ in: McDonald, G. und D. Papineau (eds.) *Teleosemantics* (Oxford: Oxford University Press), 1-22.
- Machamer, Peter, Lindley Darden und Carl F. Craver 2000: „Thinking about Mechanisms“ *Philosophy of Science* 67 (1), 1-25.
- Mahner, Martin und Mario Bunge 2000: *Philosophische Grundlagen der Biologie* (Berlin: Springer).
- Marnett, Lawrence J. 2000: „Oxyradicals and DNA damage“ *Carcinogenesis* 21, 361-370.
- Matthen, Mohan 2006: „Teleosemantics and the Consumer“ in: Macdonald, Graham und David Papineau (eds.) *Teleosemantics* (Oxford: Oxford University Press), 146-166.
- Mausfeld, Rainer 2005: „Vom Sinn in den Sinnen. Wie kann ein biologisches System Bedeutung generieren?“ in: Elsner, N. und G. Lür (eds.) „...sind eben auch nur Menschen – Verhalten zwischen Zwang, Freiheit und Verantwortung“ (Göttingen: Wallstein), 47-79.
- Maynard-Smith, John. 1992: *Evolutionsgenetik* (Stuttgart: Thieme Verlag).
- Mayr, Ernst 1961: „Cause and effect in biology“ *Science* 134, 1501–1506.
- Mayr, Ernst 1998: *Das ist Biologie: die Wissenschaft des Lebens* (Heidelberg [etc.]: Spektrum Akademischer Verlag).
- McDermott, Drew 2001: *Mind and Mechanism* (Cambridge, Mass.: MIT Press).
- McLaughlin, Peter 2001: *What Function Explain. Functional Explanation and Self-Reproducing Systems* (Cambridge: Cambridge University Press).
- Menary, Richard 2007: *Cognitive Integration. Mind and Cognition Unbounded* (Basingstoke: Palgrave Macmillan).
- Mendola, Joseph 2008: *Anti-Externalism* (Oxford: Oxford University Press).

- Melser, Derek 2004: *The Act of Thinking* (Cambridge Mass.: MIT Press).
- Millikan, R. 1984: *Language, thought, and other biological categories* (Cambridge, Mass: MIT Press).
- Millikan, R. 1986: „Thoughts Without Laws” *The Philosophical Review* 95, 47-80.
- Millikan, R. 1989a: “In Defense of Proper Functions” *Philosophy of Science* 56, 288–302.
- Millikan, R. 1989b: “Biosemantics” *Journal of Philosophy* 86, 281–97.
- Millikan, R. 1990: “Truth, Rules, Hoverflies and the Kripke-Wittgenstein Paradox” *Philosophical Review* 99, 232–53.
- Millikan, R. 1991: “Speaking Up for Darwin” in: Loewer, B. und G. Rey (eds.) *Meaning in Mind: Fodor and his critics* (Cambridge, Mass.: Blackwell), 151–165.
- Millikan, R. 1993a: “Introduction” in: dies. *White Queen Psychology and other Essays for Alice* (Cambridge, Mass.: MIT Press), 3-12.
- Millikan, R. 1993b: “Compare and Contrast Dretske, Fodor, and Millikan on Teleosemantics” in: dies. *White Queen Psychology and other Essays for Alice* (Cambridge, Mass.: MIT Press), 123-133.
- Millikan, R. 1993c: “Explanation in Biopsychology” in: dies. *White Queen Psychology and other Essays for Alice* (Cambridge, Mass.: MIT Press), 171-192.
- Millikan, R. 1993d: “Propensities, Exaptations, and the Brain” in: dies. *White Queen Psychology and other Essays for Alice* (Cambridge, Mass.: MIT Press), 31-50.
- Millikan, R. 1993e: “White Queen Psychology; or, The Last Myth of the Given” in: dies. *White Queen Psychology and other Essays for Alice* (Cambridge, Mass.: MIT Press), 31-50.
- Millikan, R. 1995: "Pushmi-pullyu-Representations" in: Tomberlin, James E. (ed.) *AI, connectionism and philosophical psychology. Philosophical perspectives* 9 (Atascadero, Calif.: Ridgeview), 185-200.
- Millikan, R. 2000: “What Has Natural Information to Do with Intentional Representation (Appendix B)” in: dies. *On Clear and Confused Ideas* (Cambridge: Cambridge University Press), 217-237.
- Millikan, R. 2002: “Biofunctions: Two Paradigms” in: Cummins, R., A. Ariew, M. Perlmann (eds.) *Functions: New Readings in the Philosophy of Psychology and Biology* (Oxford: Oxford University Press), 113-143.
- Millikan, R. 2003: “Teleological Theories of Mental Content” in: Nadel, L. (ed.) *The Encyclopedia of Cognitive Science* (New York: Macmillan).
- Millikan, R. 2008: *Die Vielfalt der Bedeutung* (Frankfurt am Main: Suhrkamp) [engl. Original: 2004: *Varieties of Meaning* (Cambridge, Mass: MIT Press)].
- Millikan, R. 2010: "On Knowing the Meaning; With a Coda on Swampman" *Mind* 119 (473), 43-81.

Millikan, R. 2013a: "Reply to Rescorla" in: Kingsbury, J. , D. Ryder und K. Williford (eds.) *Millikan and Her Critics* (Oxford: Blackwell), 103-106.

Millikan, R. 2013b: "Reply to Neander" in: Kingsbury, J. , D. Ryder und K. Williford (eds.) *Millikan and Her Critics* (Oxford: Blackwell), 37-40.

Millikan, R. 2013c: "Reply to Shea" in: Kingsbury, J. , D. Ryder und K. Williford (eds.) *Millikan and Her Critics* (Oxford: Blackwell), 81-86.

Morris, Charles 1938: „Foundations of the Theory of Signs“ in: Neurath, O., R. Carnap, C. Morris (eds.) *International Encyclopedia of Unified Science* (Chicago: The University of Chicago Press), 1-59.

Nagel, Ernst 1961: *The Structure of Science* (London: Routledge).

Nanay, B. 2010: "A Modal Theory of Content" *Journal of Philosophy* 107, 412–431.

Nanay, B. 2012: "Function Attributions Depend on the Explanatory Context: A Reply to Neander and Rosenberg's Reply to Nanay" *Journal of Philosophy* 109, 623–627.

Nanay, Bence 2014: "Teleosemantics without Etiology" *Philosophy of Science*, Vol. 81, No. 5, 798-810.

Nannini, Sandro 2004: "Mental causation and intentionality in a mind naturalising theory" in: Alberto Peruzzi (ed.) *Mind and Causality* (Amsterdam/Philadelphia: John Benjamins Publishing Company), 69-96.

Neander, Karen 1988: „What Does Natural Selection Explain? Correction to Sober“ *Philosophy of Science* Vol. 55, No. 3, 422-426.

Neander, K. 1991: "Functions as Selected Effects" *Philosophy of Science* 58, 168–184.

Neander, K. 1995: "Malfunctioning and Misrepresenting" *Philosophical Studies* 79, 109–141.

Neander, K. 2006: "Content for Cognitive Science" in: McDonald, G. and D. Papineau (eds.) *Teleosemantics* (Oxford: Oxford University Press), 167–194.

Neander, Karen 2012: "Teleological Theories of Mental Content", The Stanford Encyclopedia of Philosophy (Spring 2012 Edition), Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/spr2012/entries/content-teleological/>>.

Neander, K. 2013: "Toward an Informational Teleosemantics", in: Kingsbury, J. , D. Ryder und K. Williford (eds.) *Millikan and Her Critics* (Oxford: Blackwell), 21-36.

Neander, K. (in Vorbereitung) *The Natural and the Normative*.

Neander, K. und A. Rosenberg 2010: "Solving the Circularity Problem for Functions: A Response to Nanay" *Journal of Philosophy* 109, 613–622.

Nichols, Shaun and Stephen P. Stich 2003: *Mindreading. An Integrated Account of Pretence, Self-Awareness, and Understanding Other Minds* (Oxford: Oxford University Press).

Noe, Alva 2009: *Out of Our Heads. Why You Are Not Your Brain, and Other Lessons from the Biology of Consciousness* (New York: Hill and Wang).

Nozick, Robert 1981: *Philosophical Explanations* (Oxford: Clarendon Press).

Papineau, David 2002: *Thinking about consciousness* (Oxford : Clarendon Press).

Papineau, D. 1984: "Representation and Explanation" *Philosophy of Science* 51, 550–72.

Pellen, Luke 2008: „How not to Imitate a Human Being. An Essay on Passing the Parsing the Turing Test“ in: Robert Epstein, Gary Roberts and Grace Beber (eds.) *Parsing the Turing Test. Philosophical and Methodological Issues in the Quest for the Thinking Computer* (Springer).

Perlman, Mark 2000: *Conceptual Flux. Mental Representation, Misrepresentation, and Concept Change* (Dordrecht: Kluwer Academic Publishers).

Pierce, A. Benjamin 2007: *Genetics. A Conceptual Approach* (New York: W.H. Freeman and Company).

Pietroski, Paul M. 1992: "Intentional and Teleological Error" *Pacific Philosophical Quarterly* 73, 267–81.

Polger, Thomas W. 2004: *Natural Minds* (Cambridge, Mass.: MIT Press).

Post, John F. 1999: "naturalism" in: Robert Audi (ed.) *The Cambridge dictionary of philosophy* (Cambridge : Cambridge University Press), 596-597.

Price, C. 2001: *Functions in Mind: A Theory of Intentional Content* (Oxford: Clarendon Press).

Prinz, Jesse J. 2002: *Furnishing the Mind. Concepts and Their Perceptual Basis* (Cambridge, Mass.: MIT Press).

Prior, Elizabeth W., Robert Pargetter, Frank Jackson 1982: „Three These about Dispositions“ *American Philosophical Quarterly* Vol. 19, Number 3, 251-257.

Putnam, Hilary 1992: *Renewing Philosophy* (Cambridge, Mass.: Harvard University Press).

Rakova, Marina 2006: *Philosophy of Mind A–Z* (Edinburgh: Edinburgh University Press).

Rakover, Sam S. 2007: *To understand a cat. methodology and philosophy* (Amsterdam: John Benjamins).

Ramsey, William M. 2007: *Representation Reconsidered* (Cambridge: Cambridge University Press).

Rescorla, Michael 2013: „Millikan on Honeybee Navigation and Communication“ in: Kingsbury, J. , D. Ryder und K. Williford (eds.) *Millikan and Her Critics* (Oxford: Blackwell), 87-102.

Robinson, William S. 2004: *Understanding Phenomenal Consciousness* (Cambridge: Cambridge University Press).

Rosenberg, Gregg 2004: *A Place for Consciousness. Probing the Deep Structure of the Natural World* (Oxford: Oxford University Press).

Rothenbuhler, W.C. 1964: „Behavior genetics of nest cleaning in honey bees. IV. Responses of F1 and backcross generations to disease-killed brood“ *American Zoologist* 4, 111-123.

Rowlands, M. 2006: *Body Language. Representation in Action* (Cambridge: MIT Press).

Rupert, Robert D. 2009: *Cognitive Systems and the Extended Mind* (Oxford: Oxford University Press).

Salmon, Wesley C. 1998: *Causality and Explanation* (Oxford: Oxford University Press).

Salmon, Wesley C. 2006 [1989]: *Four Decades of Scientific Explanation* (Pittsburgh: University of Pittsburgh Press).

Samson, Birgit und Wolfgang Detel 2002: „Zum Begriff nicht-mathematischer Funktionen“ *Analyse & Kritik* 24, 100-129.

Sapire, David 1999: „disposition“ in: Robert Audi (ed.) *The Cambridge Dictionary of Philosophy* (Cambridge: Cambridge University Press), 238-239.

Schlicht, Tobias. 2008: „Ein Stufenmodell der Intentionalität“ in: Spät, P. (ed.) *Zur Zukunft der Philosophie des Geistes* (Paderborn: mentis), 59-92.

Scholz, Oliver R. 2009 [1991]: *Bild, Darstellung, Zeichen: philosophische Theorien bildhafter Darstellung* (Frankfurt, Main: Klostermann).

Schulte, Peter 2012: „How Frogs See the World: Putting Millikan's Teleosemantics to the Test“ *Philosophia* 40, 483-496.

Searle, John 2007: „Putting Consciousness Back in the Brain: Reply to Bennett and Hacker, Philosophical Foundations of Neuroscience“ in: Maxwell Bennett, Daniel Dennett, Peter Hacker and John Searle *Neuroscience and Philosophy Brain, Mind, and Language* (New York: Columbia University Press), 97-124.

Searle, John R. 2002: *Consciousness and Language* (Cambridge: Cambridge University Press).

Shannon, Claude E. und Warren Weaver 1998 [1949]: *The Mathematical Theory of Communication* (Urbana/Chicago: University of Illinois Press).

Shannon, Claude E. 1998 [1949]: „The Mathematical Theory of Communication“ in: Claude E. Shannon und Warren Weaver *The Mathematical Theory of Communication* (Urbana/Chicago: University of Illinois Press), 29-125.

Shapiro, L.A. 1992: „Darwin and Disjunction: Optimal Foraging Theory and Univocal Assignments of Content“ *Proceedings of the Philosophy of Science Association* 1, 469-480.

Shea, N. 2005: *On Millikan* (Belmont: Thomson Wadsworth).

Shea, N. 2013: „Millikan's Isomorphism Requirement“ in: Kingsbury, J. , D. Ryder und K. Williford (eds.) *Millikan and Her Critics* (Oxford: Blackwell), 63-80.

- Skyrms, Brian 2010: *Signals. evolution, learning, & information* (Oxford: Oxford University Press).
- Slezaks, Peter 2004: „The World Gone Wrong? Images, Illusions, Mistakes and Misrepresentations“ in: Clapin, H. et al. (eds.) *Representation in Mind: New Approaches to Mental Representation* (Amsterdam: Elsevier), 163-185.
- Sober, Elliott 1984a: *The Nature of selection: evolutionary theory in philosophical focus* (Cambridge - Mass.& London: The MIT Press).
- Sober, Elliott 1984b: „Force and disposition in evolutionary theory“ in: Christopher Hookway (ed.) *Minds, machines and evolution* (Cambridge: Cambridge University Press), 43-61.
- Sober, Elliott 1994: „The Primacy of Truth-Telling and the Evolution of Lying“ in: ders. (ed.) *From a Biological Point of View* (Cambridge: Cambridge University Press), 71-92.
- Sober, Elliott 2000: *Philosophy of Biology* (Boulder: Westview Press), 2nd Edition.
- Stegmann, Ulrich 2005: „John Maynard Smith’s notion of animal signals“ *Biology and Philosophy* 20, 1011-1025.
- Stegmann, Ulrich 2009: „A Consumer-Based Teleosemantics for Animal Signals“ *Philosophy of Science*, 76, 864-875.
- Stein, B. E., M.T. Wallace and T.R. Stanford 1998: „Single Neuron Electrophysiology“ in: Bechtel, W. und George Graham (eds.) *A Companion to Cognitive Science* (Malden, Mass: Blackwell), 433-449.
- Stepanians, Markus S. 2001: *Gottlob Frege zur Einführung* (Hamburg: Junius).
- Sterelny, K. 1990: *The Representational Theory of Mind: An Introduction* (Cambridge, Mass.: Blackwell).
- Stich, Stephen 1992: „What Is a Theory of Mental Representation?“ *Mind*, New Series, Vol. 101, No. 402, 243-261.
- Stout, Rowland 2006: *The Inner Life of a Rational Agent. In Defence of Philodophical Behaviorism* (Edinburgh: Edinburgh University Press).
- Strawson, Galen 2010: *Mental Reality* (Cambridge, Mass.: MIT Press).
- Szabó, Zoltán Gendler 2013: „Compositionality“, *The Stanford Encyclopedia of Philosophy* (Fall 2013 Edition), Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/fall2013/entries/compositionality/>>.
- Tetens, Holm (2006) *Philosophisches Argumentieren. Eine Einführung* (München: Beck), 2. Aufl.
- Thiel, C. 1996: „Theorie“ in: Jürgen Mittelstraß (ed.) *Enzyklopädie Philosophie und Wissenschaftstheorie* (Stuttgart/Weimar: Metzler), Bd. 4, 260-270.

Thompson, Evan 2007: *Mind in life. biology, phenomenology, and the sciences of mind* (Cambridge Mass.: Harvard University Press).

Thornton, Tim 2006: „The discursive turn, social constructionism, and dementia“ in: Hughes, Julian C., Stephen J. Louw and Steven R. Sabat (eds.) *Dementia. mind, meaning, and the person* (Oxford: Oxford University Press), 123-141.

Tinbergen, Nikolaas 1963: „On aims and methods of Ethology“ *Zeitschrift für Tierpsychologie* 20(4), 410–433.

Uhlhaas, Peter J. and Wolf Singer 2011: „Brain Evolution and Cognition: Psychosis as Evolutionary Cost for Complexity and Cognitive Abilities in Humans“ in: Welsch, Wolfgang, Wolf Singer und André Wunder (eds.) *Interdisciplinary Anthropology. Continuing Evolution of Man* (Berlin: Springer), 1-17.

Van Gelder, T. 1995: „What might cognition be, if not computation?“ *The Journal of Philosophy*, 92 (7), 345-381.

Van Gelder, T. 1998: „The dynamical hypothesis in cognitive science“ *Behavioral and Brain Sciences* 21 (5), 615-627.

Van Gelder, T. und R. Port 1995: „Its about time: An overview of the dynamical approach to cognition“ in: dies. (eds.) *Mind as motion: Explorations in the dynamics of cognition* (Cambridge, Mass.: Bradford/MIT Press).

Velmans, Max 2009: *Understanding Consciousness* (London: Routledge).

Vicari, Giuseppe 2008: *Beyond Conceptual Dualism. Ontology of Consciousness, Mental Causation, and Holism in John R. Searle's Philosophy of Mind* (Amsterdam: Rodopi).

Vogel, Matthias 2010: „Am Leben vorbei? Ruth G. Millikans Theorie der Eigenfunktionen in der Diskussion“ *Deutsche Zeitschrift für Philosophie* 58 (6), 913-934.

Walter, Henrik 1997: „Proper Functions in der Klemme? Zur Kritik teleologischer Bedeutungstheorien“ in: *Analytomen* 2 (Berlin [u.a.]: de Gruyter), 186-193.

Watkins, John 1984: *Science and Scepticism* (London: Hutchinson).

Watson, James D. und Francis H. C. Crick 1953: „Genetical Implications of the Structure of Deoxyribonucleic Acid“ *Nature* 171, 964-7.

Watson, James D. und Francis H. C. Crick 1953: „Molecular Structure of Nucleic Acids. A Structure for Deoxyribose Nucleic Acid“ *Nature* 171, 737-8.

Weaver, Warren 1998 [1949]: „Some Recent Contributions to the Mathematical Theory of Communication“ in: Shannon, Claude E. und Warren Weaver *The Mathematical Theory of Communication* (Urbana/Chicago: University of Illinois Press), 3-28.

v. Weizsäcker, Carl Friedrich 1977: *Der Garten des Menschlichen. Beiträge zur geschichtlichen Anthropologie* (München/Wien : Hanser).

Wegner, Daniel M. 2005: „Who Is the Controller of Controlled Processes?“ in: Hassin, Ran R., James S. Uleman, and John A. Bargh (eds.) *The New Unconscious* (Oxford: Oxford University Press), 19-36.

Wiener, Norbert 1948: *Cybernetics, or control and communication in the animal and the machine* (Paris: Hermann).

Wild, M. 2010a: *Biosemantik. Repräsentation, Intentionalität, Norm*. Habilitationsschrift (Berlin: ???).

Wild, M. 2010b: „Schwerpunkt: Die Philosophie von Ruth G. Millikan“ *Deutsche Zeitschrift für Philosophie* 58, 889-892.

Wiseman, Helen und Barry Halliwell 1996: „Damage to DNA by reactive oxygen and nitrogen species: role in inflammatory disease and progression to cancer“ *Biochemical Journal* 313, 17-29.

Wolpert, Lewis 1991: *The triumph of the embryo* (Oxford [etc.]: Oxford University Press).

Wunder, André 2008: *Die Semantik der genetischen Information: Eine Untersuchung ausgewählter Texte* (Saarbrücken: Vdm Verlag).

CV

PERSÖNLICHE DATEN

Vor- und Nachname: André Wunder
Geburtsdatum und -ort: 27.08.1978 in Rudolstadt, Deutschland
Nationalität: deutsch

AUSBILDUNG

Seit 09/2010 Doktoratsstudent im Fach Philosophie (Universität Zürich)
SNF-Stipendium im ProDoc „Menschliches Leben“

10/2000-09/2007 M.A. in Philosophie (Nebenfächer: Psychologie, Soziologie)
Friedrich-Schiller-Universität (FSU) Jena, Deutschland
Gesamturteil: sehr gut

09/1998-06/2000 Anwärterjahr, gehobene Beamtenlaufbahn (Forst), Studium
der Forstwirtschaft an FH Schwarzburg (abgebrochen)

1997 Abitur am Friedrich-Fröbel-Gymnasium in Bad Blankenburg,
Deutschland

ANSTELLUNGEN

Seit 01/2014 Forschungsmanager bei EU GrantsAccess (ETH Zürich und
UZH)

09/2010–08/2013 Wissenschaftlicher Mitarbeiter (SNF) am Lehrstuhl für
Theoretische Philosophie II, Philosophisches Seminar (UZH)

11/2007–12/2009 Koordinator des Forschungsverbundes des Bundes-
ministeriums für Forschung und Bildung (BMBF) -
„Interdisziplinäre Anthropologie“, (FSU Jena)

Zürich, 06.10.2015